

pour obtenir le grade de
Docteur de l'Ecole des Mines de Paris
Spécialité “ Informatique et Temps Réel, Robotique, Automatique ”

présentée et soutenue publiquement par
Claire THOMAS

le 11 décembre 2006

FUSION D'IMAGES DE RESOLUTIONS SPATIALES DIFFERENTES

Directeur de thèse : Lucien WALD

Jury :

M. Vicent Caselles	Rapporteur
M. Jocelyn Chanussot	Rapporteur
M. Philippe Réfrégier	Examineur
M. Jean-Paul Marmorat	Examineur
M. Olivier Goretta	Examineur
M. Lucien Wald	Directeur de thèse

Sommaire

REMERCIEMENTS

I	INTRODUCTION A LA FUSION D'IMAGES ET OBJECTIF DE LA THESE	1
I.1	Présentation du problème	1
I.2	Illustration des contraintes liées la synthèse d'images à haute résolution spatiale	4
I.2.1	Occultations d'objets	5
I.2.2	Inversions de contraste	8
I.3	Objectif de la thèse	9
I.4	Démarche suivie	10
I.5	Notations	12
II	QUALITE	14
II.1	Introduction	14
II.1.1	Les enjeux de la qualité des images fusionnées	14
II.1.2	Les paramètres qui influencent la qualité des produits fusionnés	15
II.2	Proposition d'un nouveau protocole d'évaluation de la qualité des produits de fusion	15
II.2.1	De la littérature aux propriétés	15
II.2.2	La propriété de cohérence	16
II.2.3	La propriété de synthèse	17
II.3	Le problème de la référence	17
II.3.1	Lacunes rencontrées dans la littérature	18
II.3.2	Proposition de référence et vérification de la propriété de synthèse	19
II.3.3	Application du protocole	20
II.3.4	Conclusion et limites du protocole	20
II.4	Les outils de l'évaluation de la qualité des produits de fusion	21
II.4.1	Analyse visuelle	22
II.4.2	Analyse quantitative	24
II.4.3	Redondances/complémentarités entre distances	29
II.4.4	Proposition de bilans de qualité	32

III	ETUDE DE L'HYPOTHESE D'EXTRAPOLATION DU BILAN DE QUALITE	34
III.1	Démarche suivie	34
III.2	Images sélectionnées pour l'étude	35
III.3	Choix des bilans de qualité (BQ)	36
III.4	Exemples de résultats obtenus par la plate-forme qualité	37
III.5	Résultats statistiques pour la vérification de l'hypothèse	40
III.5.1	Une vérification par cas	40
III.5.2	Présentation de la forme des résultats	41
III.5.3	Résultats pour les BQ monomodaux	42
III.5.4	Résultats pour les BQ multimodaux	43
III.5.5	Résultats pour les BQ globaux	43
III.6	Description et interprétation des résultats	44
III.6.1	Grandes tendances	44
III.6.2	Description des tableaux de résultats	45
III.7	Critique de la démarche de l'étude empirique	47
III.7.1	Remarque sur la distance Q	47
III.7.2	Influence de l'algorithme de décomposition multiéchelle	47
III.7.3	Identification des images défavorables	48
III.7.4	Analyse complémentaire : vérification sur des extraits Pléiades	52
III.8	Conclusion et perspectives	55
IV	UN NOUVEL OUTIL POUR L'EVALUATION DE LA QUALITE GEOMETRIQUE BASE SUR LA FTM	57
IV.1	Introduction	57
IV.1.1	Etalonnage et fonction de transfert de modulation	57
IV.1.2	Définition de la <i>FTM</i>	57
IV.2	Présentation de notre méthode d'estimation de la <i>FTM</i>	61
IV.2.1	Limite des méthodes existantes	61
IV.2.2	Principe de l'estimation de la <i>FTM</i>	61
IV.2.3	Calcul de l'équation de la droite du contour	63
IV.2.4	Du profil suréchantillonné à la <i>FTM</i>	65
IV.2.5	Validation de l'algorithme	67

IV.2.6	Conclusion sur la méthode d'estimation et perspectives d'amélioration	68
IV.3	Proposition d'une nouvelle distance pour l'évaluation de la qualité géométrique des produits de fusion (L_2CFC)	69
IV.3.1	Extrait d'une arête d'un immeuble de Toulouse	69
IV.3.2	Extrait de la prison de Toulouse	71
IV.4	Conclusion	77

V	CATEGORIES DE METHODES DE FUSION	79
----------	---	-----------

V.1	Introduction	79
V.2	Les méthodes de type projection-substitution (<i>COS</i>)	80
V.2.1	Rappel sur la notion d'espace de représentation des couleurs et principe de ce type de méthodes	80
V.2.2	La méthode <i>IHS</i> linéaire	81
V.2.3	La méthode <i>ACP</i>	82
V.3	Les méthodes de type contribution spectrale relative	83
V.3.1	Hypothèse de base de ces méthodes	83
V.3.2	La méthode Brovey	83
V.3.3	La méthode <i>CN</i> (<i>Color Normalized</i>)	83
V.3.4	La méthode <i>P+XS</i> (CNES)	83
V.4	Le concept <i>ARSIS</i>	84
V.4.1	Présentation du concept	84
V.4.2	Quelques exemples réalisations du concept	88
V.4.3	Le ratio 4 dans le cas particulier d'un <i>MSM</i> dyadique	93
V.5	Conclusion à propos des différents types de méthodes	94
V.5.1	Méthodes de type projection-substitution et de contribution spectrale relative	94
V.5.2	Méthodes de type <i>ARSIS</i>	95
V.5.3	Rappel sur les précautions à prendre concernant la qualité d'une méthode	96
V.6	Critique des méthodes de fusion récentes	96
V.6.1	La méthode de Ballester <i>et al.</i> (2003)	96
V.6.2	Méthodes combinant <i>IHS</i> à <i>ARSIS</i> ou à la contribution spectrale relative	100
V.6.3	Retour à la définition de la fusion pour définir les voies de développement prometteuses	101

VI	NOUVELLES METHODES DE FUSION	103
-----------	-------------------------------------	------------

VI.1	Artefacts des méthodes de type <i>ARSIS</i>	104
VI.1.1	Méthodes de fusion testées et illustrations choisies	104

VI.1.2	Bilan de la qualité visuelle méthode par méthode	110
VI.1.3	Description et explications des artefacts visuels	111
VI.1.4	Conclusion : vers de nouvelles méthodes de fusion	121
VI.2	Description et évaluation de nouvelles méthodes de fusion	122
VI.2.1	<i>ATWT-M3-FONC</i>	122
VI.2.2	<i>ATWT-SharpenedM3</i>	122
VI.2.3	<i>ATWT-EnhancedRWM</i>	123
VI.3	Evaluation de la qualité des nouvelles méthodes	124
VI.3.1	Evaluation visuelle des produits de fusion à haute résolution	124
VI.3.2	Evaluation des propriétés de cohérence et de synthèse	135
VI.4	Conclusion	143

VII	CONCLUSION, RECOMMANDATIONS ET PERSPECTIVES	145
------------	--	------------

BIBLIOGRAPHIE	150
----------------------	------------

LISTE DES FIGURES	158
--------------------------	------------

LISTE DES TABLEAUX	164
---------------------------	------------

ANNEXES	
----------------	--

I Introduction à la fusion d'images et objectif de la thèse

1.1 Présentation du problème

Un des éléments clefs pour l'observation de la Terre et la connaissance de notre environnement est formé par les systèmes spatiaux. Ces satellites emportent à leur bord divers capteurs actifs (radars) ou passifs. Les capteurs qui nous intéressent sont les capteurs passifs dont la bande passante est située dans le visible et l'infrarouge. Ils peuvent présenter des caractéristiques très différentes. Les images délivrées par ces capteurs permettent la distinction des structures géométriques en fonction de la résolution spatiale que l'on considérera ici comme égale au pas d'échantillonnage correspondant à la taille du pixel. Certains capteurs sont capables d'intégrer l'énergie radiative incidente à l'entrée du capteur sur une large bande en longueur d'onde, et offrent donc peu d'information au niveau spectral, mais proposent en même temps une résolution spatiale élevée. En général, ce type d'images s'appelle « image panchromatique », noté Pan ci-après. Au contraire, d'autres capteurs embarqués sur ces mêmes satellites travaillent sur des bandes beaucoup plus étroites du spectre ; ce sont les images multispectrales, notées MS, ci-après ou hyperspectrales dans le cas où le nombre de bandes pour un même capteur dépasse la vingtaine. Leur résolution spectrale est beaucoup plus élevée, mais c'est au prix d'une faible résolution spatiale. Notons que les termes de résolution 'plus élevée' ou 'plus faible' sont des termes relatifs pour qualifier les résolutions des différentes images entre elles, acquises par un même système d'observation.

Pour de nombreuses applications, l'avantage de manipuler des images multispectrales possédant une très haute résolution spatiale a déjà été démontré, par exemple pour la classification (Gonzalez *et al.* 1996, Tsagaris *et al.* 2005, Huang *et al.* 2003, Byeungwoo et Landgrebe 1999), la détection de cibles (Fillipidis *et al.* 2000), la cartographie de la voirie (Couloignier *et al.* 1998), l'étude du développement des zones urbaines (Zhang et Wang 2002, Sengupta *et al.* 2004), la prospection archéologique (Ricchetti 2004) ou encore dans le domaine médical pour l'aide au diagnostic (Constantinos *et al.* 2001). Il est bien sûr possible de distinguer de larges zones dans une image MS ou Pan dès lors qu'elles possèdent des caractéristiques très distinctes, comme la surface de l'océan par rapport à la zone côtière, ou une exploitation agricole par rapport à une région très urbanisée. Par contre, il devient nettement plus difficile de distinguer des espaces végétaux ou minéraux très hétérogènes, sans considérer leurs signatures spectrales. L'information de 'couleurs' favorise la distinction des différentes régions de l'image grâce à des nuances jusqu'alors invisibles dans l'image Pan initiale en niveaux de gris. Par conséquent, si l'on peut utiliser ces contenus spectraux tout en profitant d'une meilleure résolution spatiale, l'identification des objets d'une scène sera enrichie et plus précise (Cornet *et al.* 2004).

Il est alors légitime de se demander pourquoi les constructeurs n'ont pas directement construit des capteurs capables de délivrer des images à hautes résolutions spatiale et spectrale. Ces notions sont contradictoires d'un point de vue technique ; plus la bande spectrale d'acquisition d'un capteur est large, plus le flux lumineux incident sur un pixel de l'image est grand. L'ouverture de la pupille d'entrée peut alors être réduite, la taille des pixels est alors plus faible et donc l'image présente une meilleure résolution spatiale. Par contre, une bande beaucoup plus étroite du spectre limite le nombre de photons absorbés par le capteur généralement CCD pour *Charged-Coupled Device*, qui se traduit en français par DTC pour Dispositif à Transfert de Charge (des travaux actuels permettent d'espérer passer à des technologies CMOS moins spécifiques) et donc l'énergie incidente sur chaque cellule. Le

temps d'intégration doit donc être plus grand que précédemment pour pouvoir délivrer un rapport signal sur bruit et une dynamique similaires. Comme le vecteur spatial bouge, cette augmentation du temps d'intégration entraîne une augmentation de la taille du pixel, expliquant la résolution spatiale plus faible.

Pour réduire le temps d'intégration, une solution serait l'augmentation de la sensibilité des détecteurs. Mais la limitation est tout d'abord financière, puisque construire des capteurs plus sensibles coûte cher. Outre la limitation budgétaire, d'autres contraintes sont imposées par la plate-forme satellite en terme d'espace, de poids, de puissance, de rayonnement électromagnétique, mais aussi des objectifs de mission (orbite), de la fusée pour le lancement, etc. En admettant que les capteurs soient en mesure de délivrer des images MS de résolution spatiale atteignant celle du Pan, deux autres problèmes vont se poser ; la capacité de stockage sur la plate-forme et la bande passante du canal de transmission sont limitées. Prenons l'exemple du satellite Ikonos ou Quickbird : ces satellites possèdent quatre bandes MS. Si la résolution spatiale de ces images était augmentée d'un facteur 4 pour rejoindre celle du Pan, leur taille serait multipliée par 16, soit 16 fois plus de données à stocker et à transmettre au sol.

Pour une même zone géographique, l'utilisateur se trouve donc en possession d'un jeu d'images dont certaines bénéficient d'une bonne résolution spatiale, alors que d'autres lui apportent une bonne résolution spectrale. Le tableau 1.1 synthétise les différentes informations concernant l'imagerie des satellites les plus couramment utilisés en environnement et en cartographie : SPOT4, SPOT5, Ikonos et Quickbird. Par exemple, le satellite Ikonos propose des images acquises par cinq modalités (ou capteurs) différentes. Quatre d'entre elles sont des modalités MS, avec des bandes spectrales situées dans le bleu, le vert, le rouge et le proche infrarouge (PIR). Leur résolution spatiale est de 4 m. La modalité Pan a l'avantage de proposer une image de résolution spatiale quatre fois plus importante (1 m) mais sa bande spectrale est plus large en s'étendant du bleu jusqu'au PIR inclus.

Caractéristiques des images satellites						
	Modalités MS				Modalité Pan	
satellite	nom de la bande	couleur	bande spectrale (µm)	résolution au sol (m)	bande spectrale (µm)	résolution au sol (m)
SPOT4	B1	vert	0,50 - 0,69	20	B2 : 0,61 - 0,68	10
	B2	jaune	0,61 - 0,68	10		
	B3	PIR	0,78 - 0,89	20		
	B4	MIR	1,58 - 1,75	20		
SPOT5	B1	vert	0,50 - 0,59	10	0,48 - 0,71	2,5 ou 5
	B2	jaune	0,61 - 0,68	10		
	B3	PIR	0,78 - 0,89	10		
	B4	MIR	1,58 - 1,75	10		
Ikonos	B1	bleu	0,45 - 0,53	4	0,45 - 0,90	1
	B2	vert	0,52 - 0,61	4		
	B3	rouge	0,64 - 0,72	4		
	B4	PIR	0,77 - 0,88	4		
Quickbird	B1	bleu	0,45 - 0,52	2,8	0,45 - 0,90	0,7
	B2	vert	0,52 - 0,60	2,8		
	B3	rouge	0,63 - 0,69	2,8		
	B4	PIR	0,76 - 0,90	2,8		

Tableau 1.1 : exemples de systèmes d'observation de la Terre offrant diverses images à différentes résolutions spatiales.

Comme la haute résolution spectrale et la haute résolution spatiale sont contenues dans des images différentes, le problème devient celui d'une synthèse d'images multispectrales possédant la meilleure résolution spatiale disponible au sein du jeu de données. La fusion d'images est un moyen de proposer une exploitation conjointe des informations provenant à l'origine de sources distinctes. D'après Wald (1999), il s'agit d'un cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance des données provenant de sources diverses. La fusion vise à l'obtention d'information de plus grande qualité, celle-ci dépendant de l'application et de son utilisateur.

Le terme de fusion englobe plusieurs définitions possibles. Par exemple, les travaux de Piella (2003), Redondo *et al.* (2005), Chen et Varshney (2005) et Zheng *et al.* (2005) considèrent la fusion comme une combinaison d'informations pertinentes des deux images placées en entrée, l'information dite « pertinente » dépendant de l'application. Par exemple, certains de ces auteurs montrent des exemples de photos ou d'images acquises par un même capteur mais à des focales différentes. Ils définissent la fusion comme la synthèse d'une image qui serait focalisée en tout point, c'est-à-dire présentant la plus haute résolution spatiale des deux images pour chaque pixel. Cette idée de complémentarité se retrouve aussi en imagerie médicale puisque des images radio, obtenues par résonance magnétique nucléaire ou encore par tomographie par émission de positons, délivrent des informations très différentes sur le corps humain (Profanter *et al.* 2004). La fusion vise à superposer les informations pertinentes provenant de différentes sources pour l'aide au diagnostic. Cette manière de définir la fusion d'images a la particularité de ne pas tenir compte des caractéristiques d'un capteur en particulier. L'enjeu se situe dans la création d'une image hybride présentant des caractéristiques spécifiques sélectionnées dans les différents ensembles placés en entrée.

Le cadre de ce travail de recherche que nous avons adopté est le cas particulier de la synthèse d'images multimodales à haute résolution spatiale : chaque image synthétisée doit être aussi proche que possible de celle qu'aurait pu observer le capteur correspondant s'il possédait cette résolution spatiale. Ce dernier énoncé constituera pour nous la propriété fondamentale des produits de fusion. L'intérêt économique de créer des images reproduisant ce que devrait observer un capteur multispectral à meilleure résolution spatiale est réel dans le domaine spatial, puisqu'une méthode de fusion performante permet de se contenter de capteurs plus simples et donc moins coûteux, et plus légers, d'où un gain de poids et d'espace sur la plate-forme.

Notons que le formalisme mathématique des méthodes de fusion, à l'origine, et dans ce texte, développé pour des images satellitales à haute résolution spatiale, permet d'étendre le champ d'application d'une méthode à d'autres objets. Par exemple, Cros (2004) a utilisé certains algorithmes de fusion décrits ci-après pour obtenir des produits de fusion représentant des cartes climatiques de propriétés optiques des aérosols atmosphériques à plus haute résolution spatiale.

1.2 Illustration des contraintes liées à la synthèse d'images à haute résolution spatiale

La figure 1.1 illustre la fusion d'images telle que nous l'entendons. L'image de gauche correspond à la composition colorée des bandes rouge, verte et bleue du satellite Ikonos. La scène est un extrait de Blagnac, aéroport international de Toulouse. Sa résolution spatiale est de 4 m. Nous l'avons suréchantillonnée par un opérateur bicubique. L'image de droite représente la composition colorée des modalités correspondantes fusionnées à la résolution du Pan (1 m). Sur l'image originale, nous ne sommes pas capables de compter le nombre de véhicules sur le parking de l'aéroport ou la signalisation sur le tarmac. Dans l'image de synthèse, les contours ont été affinés et des petits objets sont apparus. La résolution spatiale permet dorénavant de visualiser les voitures et les aspérités des toits des bâtiments.

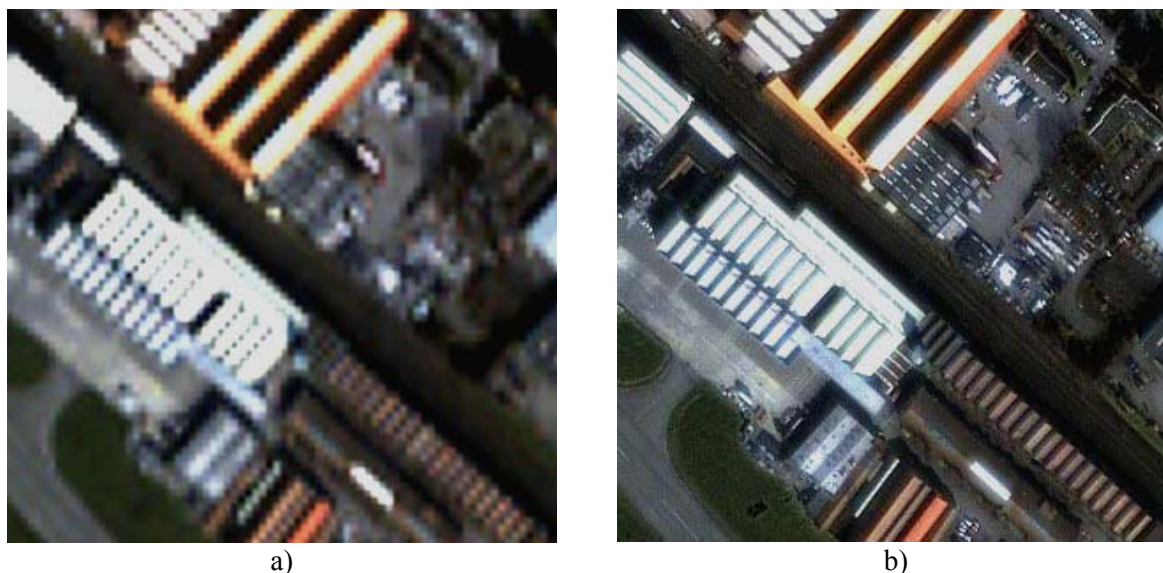


Figure 1.1 : image Ikonos, partie de l'aéroport (Blagnac) de Toulouse. a) Composition colorée initiale des modalités visibles, 4 m de résolution spatiale, suréchantillonnées par un opérateur bicubique, b) composition colorée des images fusionnées à 1 m de résolution. Copyright Space Imaging 2002.

On remarque la richesse d'interprétation apportée par l'accroissement de résolution. Généralement, la modalité MS à fusionner et celle qui comporte les détails de haute résolution spatiale représentant la même zone géographique possèdent des caractéristiques proches en terme de structures géométriques, comme c'est le cas sur ces images. Ici, aucun problème flagrant n'apparaît, mais il se peut que les deux jeux d'images placés en entrée possèdent localement des caractéristiques dissemblables se manifestant par des luminances différentes. Cette notion n'est pas évidente à

maîtriser et certains auteurs rencontrent des difficultés à formuler les enjeux qui entourent la fusion d'images MS à haute résolution spatiale (Zhou *et al.* 1998, Zhang 1999). Nous avons souhaité illustrer deux types de dissimilitudes locales : les occultations d'objets et les inversions de contraste.

1.2.1 Occultations d'objets

1.2.1.1 Cas d'images acquises à des dates différentes

L'occultation peut se produire lors de la fusion entre modalités acquises à des moments différents. Si des images parfaitement recalées géométriquement ne sont pas acquises simultanément, plusieurs transformations peuvent se produire durant ce laps de temps : variations dans la végétation en cas de saisons différentes, conditions d'illumination en fonction de l'heure d'acquisition, développement de zones périurbaines ou encore disparition d'habitations suite à une inondation ... La figure 1.2 nous permet de visualiser les effets d'une construction humaine, ici d'un barrage. Il s'agit du projet de construction du barrage des Trois Gorges en Chine. La figure 1.2a correspond à la modalité Pan de SPOT2 acquise en 1990 à 10 m de résolution spatiale. La figure b montre la même scène huit ans plus tard dans la modalité verte. Les travaux du barrage ont débuté entre ces deux dates engendrant des répercussions sur la faune et la flore locale jusqu'à 700 km en amont. Les berges de la rivière ont été dégagées de toute végétation pour accueillir les travaux de construction. Les changements qui se sont produits dans les huit années séparant ces deux images, sont parfaitement perceptibles.

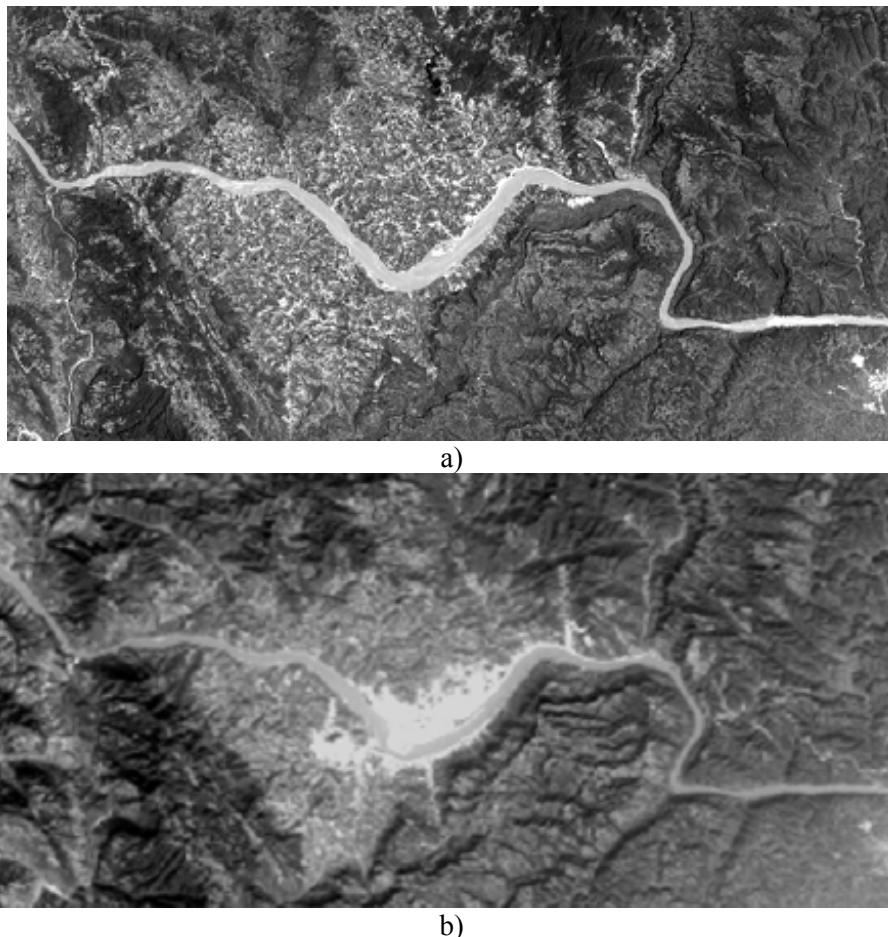


Figure 1.2 : image SPOT2, barrage des Trois Gorges (Chine). a) Modalité Pan, acquise en 1990, b) modalité verte XS1, 1998. Copyright CNES SPOT-Image 1990 et 1998.

Le comportement spectral local de deux images acquises à deux instants distincts est difficilement prévisible. Des comportements différents peuvent aussi apparaître dans des images parfaitement superposables et acquises simultanément, simplement à cause de la bande passante des différentes modalités qui vont interagir dans le processus de fusion.

1.2.1.2 Objets occultés à cause de la bande spectrale des capteurs

La figure 1.3 présente la réponse normalisée des bandes spectrales des capteurs du satellite Quickbird en fonction de la longueur d'onde en μm . La courbe noire représente la réponse spectrale de la modalité Pan de Quickbird. Les courbes de couleurs correspondent aux modalités MS, où les couleurs bleue, verte, rouge et marron correspondent respectivement aux bandes bleue, verte, rouge et PIR du capteur.

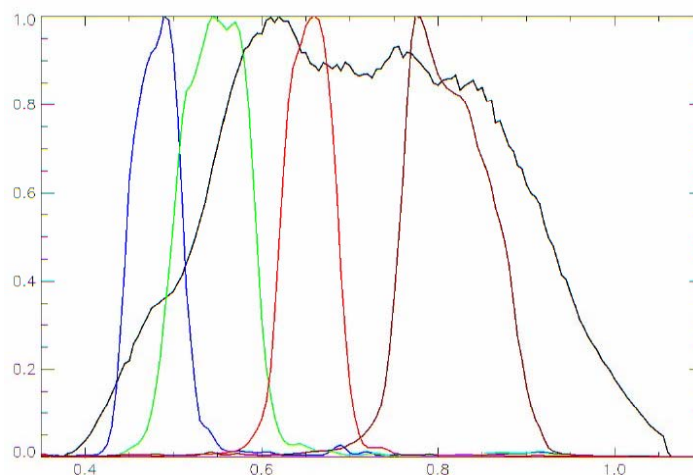


Figure 1.3 : réponses normalisées des bandes spectrales du satellite Quickbird. En abscisse : longueurs d'ondes en μm .

Un pixel d'une image quelconque présente un niveau de gris non nul à partir du moment où l'objet contenu dans ce pixel réfléchit ou émet un rayonnement dans des longueurs d'onde appartenant à la bande passante du capteur de l'image. En partant de ces considérations, la fusion d'images provenant de capteurs couvrant des bandes spectrales différentes, se heurte à la possibilité qu'un objet soit présent dans une modalité et absent d'une autre. Par exemple, si un objet réfléchit un rayonnement situé dans le moyen infrarouge dans des longueurs d'onde situées entre 0.95 et 1.05 μm alors, d'après la figure 1.3, seule la courbe noire couvre cette partie du spectre, ce qui signifie que seule la modalité Pan présentera une radiométrie qui ne sera pas égale à zéro pour cet objet.

La figure 1.4 montre des extraits d'une image Quickbird dans la modalité Pan (figure 1.4a) et bleue (figure 1.4b) d'une zone située en périphérie de la ville de Fredericton (Canada), centrée sur un terrain ne comportant pas d'habitation. La figure 1.4c représente deux sections d'une ligne localisée au centre de l'image, matérialisée par une flèche rouge dans les deux images a et b. Les modalités Pan et MS correspondent respectivement aux courbes en trait plein et en pointillés.

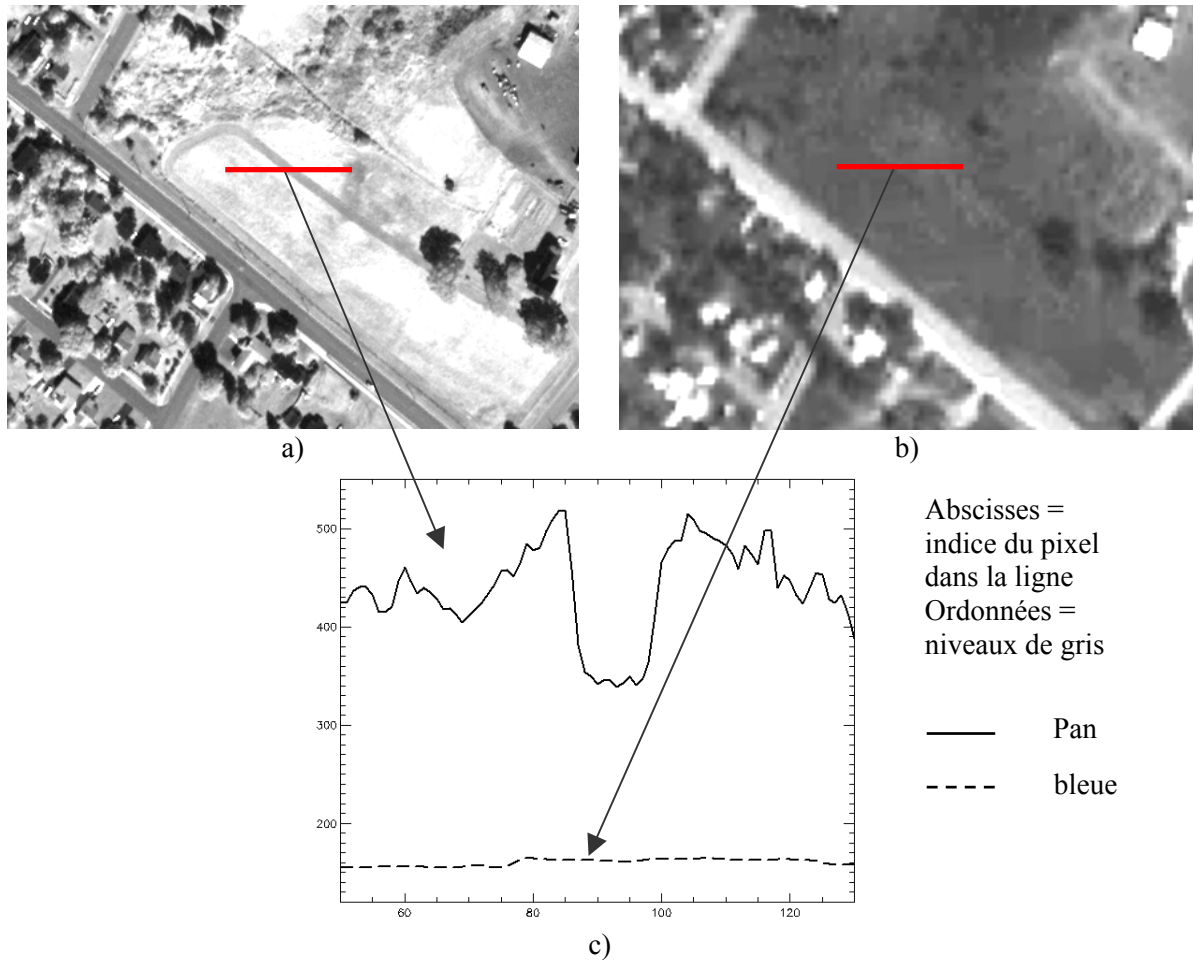


Figure 1.4 : image Quickbird, Fredericton (Canada) - illustration du phénomène d'occultation. a) Pan, 1 m, b) bleue, 4 m suréchantillonnée à 1 m (bicubique), c) extraits d'une ligne particulière de chaque image : le chemin d'herbe est absent dans l'image bleue. Copyright Digital Globe 2002.

Le chemin bien visible dans la modalité Pan est absent de la modalité bleue. Localement, cette structure est manquante dans la modalité à basse résolution spatiale et crée donc une dissimilarité entre ces deux images.

Un autre exemple est la chaussée du Bazacle, la chute d'eau de la Garonne située en plein cœur de la ville rose. La figure 1.5 montre le centre ville de Toulouse acquise par le satellite SPOT2. Le cercle blanc présent sur chaque image attire l'attention du lecteur sur la localisation de cette chute d'eau. La modalité Pan à 10 m de résolution spatiale située à gauche sur la figure, permet d'observer l'écume produite par la chute d'eau. L'écume occupe plusieurs pixels en aval, soit plusieurs dizaines de mètres. Par contre, la chaussée du Bazacle est entièrement occultée dans l'image de droite, modalité verte de ce même satellite, dont la résolution est de 20 m.

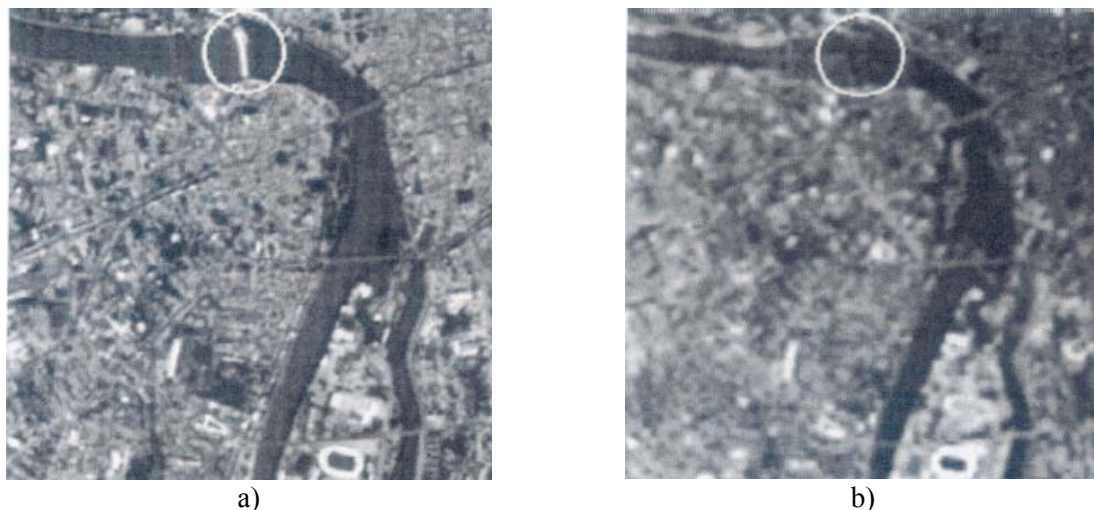


Figure 1.5 : image SPOT2, Toulouse (France). a) modalité Pan, (10 m de résolution), b) modalité XS1, 20 m rééchantillonnée (bicubique). La chute d'eau n'apparaît pas sur cette modalité. Copyright CNES 1990.

L'occultation s'explique par le fait que, dans cette modalité, aucune distinction ne peut être effectuée entre l'écume de la chute et une eau beaucoup plus calme en amont ou en aval de la chute dans des longueurs d'onde situées dans le spectre du vert. Un canal comprenant le PIR permet d'effectuer cette distinction à cause de sa forte sensibilité aux particules en suspension dans l'eau. C'est le cas de la modalité Pan dont la bande spectrale comporte une partie de PIR.

1.2.2 Inversions de contraste

Comme un objet peut être présent dans une modalité et absent d'une autre, il peut en être de même de son environnement. Ainsi, un objet clair sur fond sombre dans la modalité Pan, peut se retrouver foncé sur fond clair dans une des modalités MS, et inversement. La figure 1.6 illustre cette observation en présentant un agrandissement d'une des zones résidentielles de Fredericton et suit la même organisation que la figure 1.4. La ligne extraite couvre une région de l'image où la radiométrie du terrain change selon la modalité observée. L'absence d'ombre portée nous pousse à croire qu'il s'agit de deux terrains de nature différente, et non d'un bâtiment. D'après la radiométrie de la route et de la végétation environnante, il pourrait s'agir d'une parcelle de terrain goudronné, entourée de sol nu ou d'herbe rase.

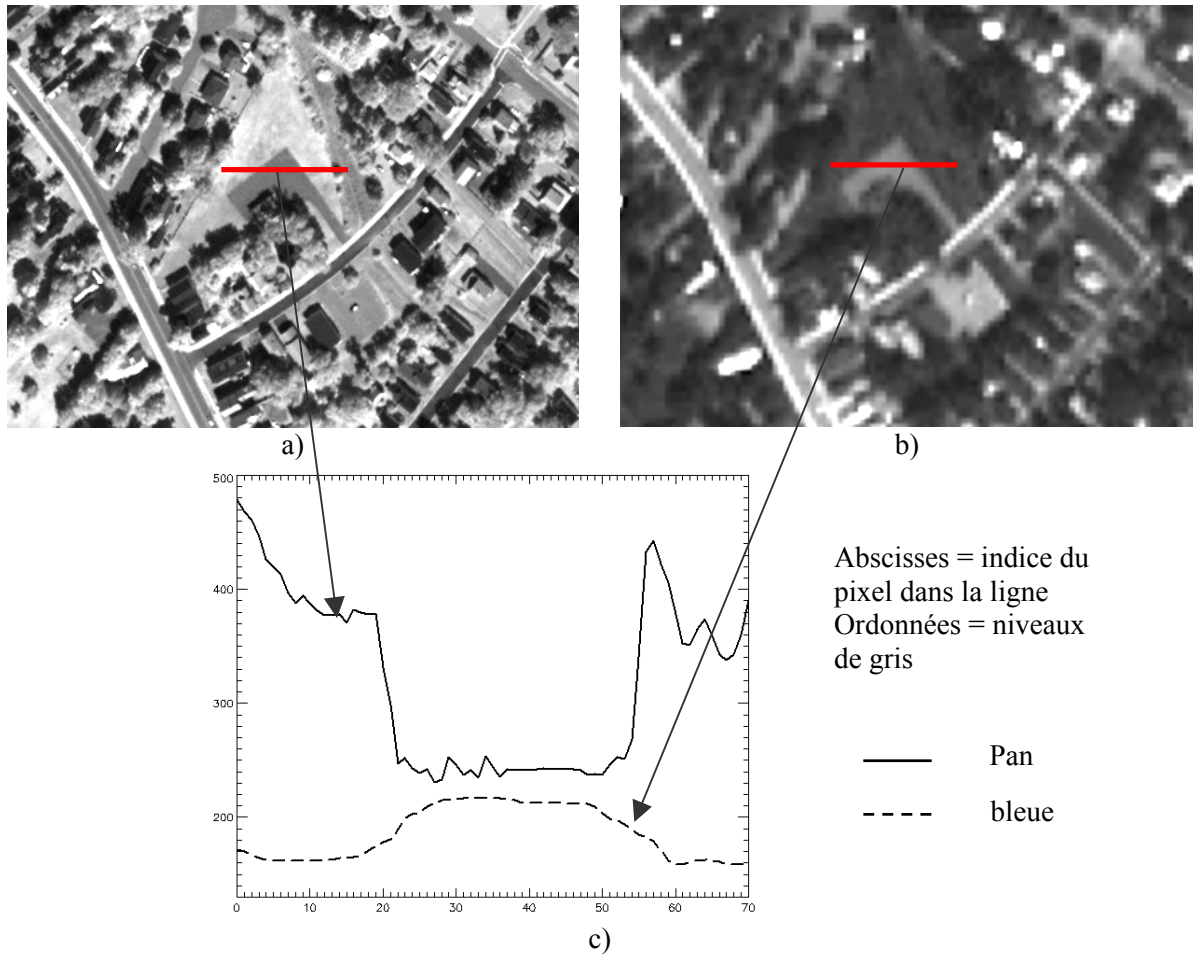


Figure 1.6 : illustration du phénomène d'inversion de contraste – image Quickbird, Fredericton (Canada). a) Pan, b) bleue, c) extraits d'une ligne particulière : le bâtiment présente un contraste inversé par rapport à son environnement dans les deux images. Copyright Digital Globe 2002.

La courbe en trait plein, qui représente les niveaux de gris de l'image Pan le long de cette ligne, montre que la parcelle rectangulaire apparaît plus sombre comparée à son environnement local. La modalité bleue expose le phénomène contraire puisque la parcelle est plus claire que la région qui l'entoure. On observe ici une inversion de contraste dans les transitions, entre la modalité à haute résolution spatiale et la modalité à fusionner.

Ces exemples montrent l'existence de dissimilitudes locales existant entre les capteurs de haute et de basse résolution qui peuvent poser problème lors de la fusion. Une méthode de fusion performante respectera les caractéristiques qui définissent le contenu de la bande spectrale à synthétiser tout en augmentant la résolution de l'image. La résolution spatiale atteinte grâce à la fusion est supposée être celle de la modalité à haute résolution spatiale ayant servi à la fusion. Tous les algorithmes de fusion (ou de synthèse) ne sont pas identiques en termes de performance, tant au niveau géométrique que radiométrique. Notons qu'il n'est pas aisé d'évaluer la qualité d'un produit de fusion puisque nous ne connaissons pas le produit idéal. En effet, aucune référence n'est disponible pour établir un bilan absolu de qualité.

1.3 Objectif de la thèse

L'objectif de la thèse est de contribuer à la compréhension du problème de la synthèse d'images à haute résolution spatiale ainsi qu'à l'évaluation de la qualité des produits fusionnés. À terme, le but est le développement d'une méthode optimisée, plus performante que les méthodes actuelles et qui respecte la propriété fondamentale des produits de fusion, ou tout au moins proposer

des pistes de développement. Ces recherches se sont accompagnées d'une formalisation mathématique autour de l'analyse multiéchelle ou multirésolution avec des outils relevant essentiellement du filtrage numérique, et autour des approches variationnelles.

La première étape du travail a consisté à comprendre les lacunes de la littérature concernant l'évaluation de la qualité dans le but d'élaborer un protocole d'évaluation de la qualité des images fusionnées générique. Il s'agissait aussi d'étudier les outils de l'évaluation en termes de distances et de bilans de qualité afin de disposer d'une bonne base d'évaluation pour nos développements. Nous avons proposé une distance pour caractériser la qualité géométrique des produits fusionnés basée sur la fonction de transfert de modulation du capteur directement estimée à partir des images acquises par celui-ci.

La seconde étape a été dédiée à l'évaluation critique des méthodes de fusion existantes et particulièrement aux méthodes les plus récentes, en s'intéressant aux voies de développement choisies par les différents auteurs. L'évaluation qualitative et quantitative de leurs performances a ouvert de nombreuses perspectives de développement, en justifiant l'importance de l'utilisation de l'analyse multiéchelle ou multirésolution. Le cadre de développement de nouvelles méthodes que nous avons choisies est celui du concept *ARSIS*. Les améliorations des trois nouvelles méthodes développées concernent la qualité géométrique de ce type de méthodes, souvent controversée par le passé. Une importance notoire a été accordée à l'identification et la classification des artefacts des méthodes de fusion de type *ARSIS*.

Enfin, un objectif technique a accompagné ce travail de thèse : la réalisation d'une plate-forme logicielle qui supporte le développement de méthodes ainsi que l'évaluation de la qualité. Le langage choisi est l'IDL (*Interactive Data Language*) qui permet facilement de traiter et manipuler les images, associé au logiciel de traitements et de visualisation d'images ENVI. Pour répondre aux impératifs industriels, nous avons développé une interface graphique sous ENVI qui propose deux sous-menus. Le premier offre l'opportunité de tester diverses stratégies de fusion de type *ARSIS*. Ces stratégies regroupent à la fois des implantations existantes de la littérature, mais aussi nos nouvelles implantations. Le second sous-menu met à disposition de l'utilisateur un module complet pour l'évaluation de la qualité.

1.4 Démarche suivie

Le développement de nouveaux algorithmes passe immanquablement par une étape d'estimation de la qualité des produits de fusion obtenus. La bibliographie a montré qu'un cadre normatif pour l'évaluation de la qualité des produits de fusion était nécessaire. En effet, aucun protocole n'était unanimement accepté et utilisé par la communauté de la fusion d'images pour évaluer la qualité de ces méthodes ou des produits issus de la fusion. Le chapitre 2 présente une étude critique des protocoles présents dans la littérature puis propose deux propriétés à vérifier qui découlent de la propriété fondamentale des produits de fusion. Le protocole que nous proposons a été discuté et accepté par la communauté de la fusion lors d'une réunion du groupe de travail de l'association européenne EARSeL (Thomas et Wald 2004). Il est applicable quelles que soient les images, le ratio ou le paysage concerné. Ce protocole s'accompagne d'une proposition de distances pour la vérification des propriétés. On montre que certaines de ces distances sont redondantes, alors que d'autres se complètent (Thomas et Wald 2006c) afin d'aboutir à une série de distances que nous appelons bilan de qualité (BQ) qui permet d'obtenir une idée relativement complète de la qualité d'un produit de fusion.

Cependant, une des hypothèses d'application du protocole, appelée « hypothèse d'extrapolation » peut être controversée. Pour connaître la qualité d'un produit de fusion, il faut le comparer à une référence, qui est manquante dans le cas de la synthèse d'images MS à haute résolution spatiale. L'hypothèse d'extrapolation consiste à appliquer le processus de fusion à une échelle où la référence est disponible et à avancer que le bilan de qualité à cette basse résolution est

proche, voire meilleur, que celui qui aurait pu être dressé à une échelle plus fine (de plus haute résolution spatiale). Nous avons entrepris la vérification empirique de cette hypothèse puisqu'aucune preuve théorique ne peut être apportée. Le mode opératoire est décrit et critiqué, de même que l'ensemble des images traitées et des distances retenues pour cette étude. Les résultats ont permis de confirmer que l'hypothèse d'extrapolation est généralement valide et de donner des informations sur son domaine de validité. Ainsi, l'usage du protocole dans notre travail se trouve justifié.

La résolution spatiale des images et leur qualité géométrique sont intrinsèquement reliées à la qualité d'un capteur d'images satellite. Le chapitre 4 est consacré à la fonction de transfert de modulation (*FTM*), grandeur physique déterminante dans la caractérisation des capteurs. Elle est rarement donnée de manière explicite par les constructeurs du capteur car elle témoigne directement des faiblesses de l'instrument. De plus, certaines images peuvent subir des post-traitements suite à leur acquisition, qui sont complètement opaques pour l'utilisateur et font de chaque image une entité unique. Voilà pourquoi nous avons construit une méthode de lecture (ou d'estimation) de la *FTM* directement à partir des images. La première partie est consacrée à la présentation de cet algorithme de lecture, à son implantation en IDL et à une étude critique des avantages et inconvénients d'une telle approche. La deuxième partie transforme la vision de cet outil d'estimation afin de l'utiliser en tant que nouvelle distance pour l'évaluation de la qualité géométrique des produits de fusion. La connaissance a priori de l'influence de plusieurs méthodes de fusion sur la qualité des contours d'une image a permis de valider cet outil.

Le chapitre 5 présente un état de l'art critique des méthodes de fusion. Plusieurs classifications des méthodes ont été proposées, et nous avons choisi celle de Wald (2002) organisée en trois grandes catégories : projection-substitution, contribution spectrale relative et enfin les méthodes qui rentrent dans le concept *ARSIS* (Amélioration de la Résolution Spatiale par Injection de Structures). *ARSIS* repose sur le principe que les informations manquantes des modalités MS sont essentiellement localisées dans les fréquences élevées du spectre des fréquences spatiales et considère que ces informations peuvent être inférées à partir des informations disponibles à l'origine, c'est-à-dire des deux ensembles Pan et MS. Une présentation des implantations principales de ces trois catégories est développée. Ensuite, on s'intéresse aux méthodes de fusion les plus récentes, en étudiant les solutions que les auteurs ont explorées pour améliorer la qualité spatiale et/ou spectrale des produits fusionnés. Il s'agit généralement de méthodes de fusion hybrides entre deux catégories de fusion, comme par exemple celle de Ballester *et al.* (2003). Le chapitre est complété par une démonstration qui montre que, dans le cadre de la synthèse de haute qualité des signatures spectrales à haute résolution spatiale, seules les méthodes utilisant une approche multirésolution ou multiéchelle sont valides. Cette démonstration justifie notre choix de développer dans le concept *ARSIS*.

Pour apporter des améliorations aux méthodes de type *ARSIS*, nous nous sommes intéressés à leur point faible, qui est la qualité géométrique. La qualité des produits fusionnés a souvent été jugée insuffisante à cause de nombreux artefacts géométriques, qui diffèrent selon l'implantation et la scène considérées. Le chapitre 6 débute par l'identification et la classification des artefacts de cinq méthodes de type *ARSIS*. La routine « fusion » de la plate-forme logicielle sous ENVI nous a permis de tester quelques-unes des implantations existantes du concept. Nous avons eu la chance de bénéficier des remarques et conclusion d'un groupe d'interprètes d'images du ministère de la défense. La première conclusion est que les produits fusionnés par quelques-unes de ces méthodes existantes représentent un apport indéniable à l'analyse des scènes. Cependant, aucun algorithme n'a été jugé suffisamment satisfaisant et fiable pour être utilisé comme tel dans une démarche opérationnelle. Les observations des interprètes et leur classification des méthodes de fusion ont été en accord avec les résultats produits par notre outil d'évaluation de la qualité. L'explication mathématique sur l'origine des défauts géométriques associée à l'analyse des bons résultats statistiques de ces méthodes nous a amenés à développer de nouveaux algorithmes qui devaient pallier au moins partiellement les artefacts identifiés. Pour cela, on montre que la nouvelle méthode de fusion doit être en mesure de proposer un équilibre entre une approche locale et globale d'estimation des paramètres. Trois nouvelles méthodes de fusion ont été développées dans cette optique. Le protocole d'évaluation de la qualité des produits de fusion a été appliqué à une nouvelle série d'images traitées par ces nouveaux algorithmes ainsi que

par des méthodes existantes, dont la méthode de Ballester *et al.* (2003). Pour l'évaluation visuelle à haute résolution spatiale qui correspond à la première étape de notre protocole, nous avons pu avoir la chance de travailler avec les mêmes interprètes d'images. Les analyses des interprètes combinées avec les résultats issus de l'application du protocole ont mené à la conclusion que la nouvelle méthode de fusion intitulée *ATWT-SharpnedM3* est la meilleure méthode de fusion de toutes celles testées dans ce chapitre de thèse. La conclusion la plus importante est qu'elle a été déclarée par les interprètes la seule à être utilisable en l'état dans un contexte d'exploitation opérationnelle au sein du ministère de la défense et pour des applications de cartographie en général.

Enfin, le chapitre 7 dresse une conclusion sur l'ensemble des travaux menés tout au long de cette thèse. Nous récapitulons les pistes explorées et expliquons pourquoi certaines d'entre elles n'ont pas abouti. Des perspectives de développement et de nombreuses recommandations sont proposées.

1.5 Notations

Le vocabulaire employé dans la littérature pour désigner les images satellites est très divers : image, canal, bande spectrale, modalité ... Le terme le plus général est « modalité » car il désigne en plus de l'image, son moyen d'acquisition. Pour la vision humaine, une modalité serait par exemple l'œil gauche.

Définissons l'ensemble des modalités de haute résolution spatiale par A_0 et l'ensemble des modalités de basse résolution par B_l . Les termes de 'haute/basse' résolution spatiale sont des termes relatifs pour désigner l'écart entre les résolutions spatiales des ensembles originaux. Ici, la haute résolution spatiale est notée $res0$. En indice, le 0 indique la haute résolution spatiale, par exemple, A_0 . La basse résolution spatiale est notée $res1$, avec 1 comme indice correspondant, par exemple B_1 . La fusion va permettre de synthétiser les modalités B_0 (tableau 1.2).

Chaque ensemble d'images A et B est composé d'une ou de plusieurs images acquises selon différentes modalités. On se place dans le cas où les images ont subi un traitement géométrique préalable pour être parfaitement superposables pixel à pixel. Dans l'ensemble B , l'image B_k désigne l'image acquise dans la bande spectrale (ou modalité) k . Par souci de simplicité, la plupart du temps, l'ensemble A sera réduit à une seule image que nous assimilerons à l'image Pan, à moins qu'il en soit spécifié autrement. Le problème pourra facilement être étendu au cas où A comprendrait plus d'une image ou dans le cas de plus de deux jeux d'images disponibles.

<i>Introduction des notations</i>	<i>Quickbird (exemple)</i>
$0, 1, 2, 3 \dots$: indices des résolutions spatiales.	
$res0, res1, res2, \dots$: résolutions spatiales successives.	$res0 = 0.7 \text{ m}, res1 = 2.8 \text{ m}, res2 = 11.2 \text{ m}, \dots$
$ratio$: rapport entre deux résolutions spatiales successives, avec $ratio = res1/res0 = res2/res1 \dots$	$ratio = 4$
N : le nombre total de modalités MS.	$N = 4$
A_0 : modalité(s) de haute résolution spatiale.	$A_0 = \text{Pan}$ et $res0 = 0.7 \text{ m}$
$(A_l)^{interp}_0$: modalité A à la résolution $res1$ rééchantillonnée pour s'adapter à la taille de A_0 .	$(A_l)^{interp}_0 = \text{Pan}$ à 2.8 m ramenée à la taille (H_1, H_2) de A_0 .
B_{kl} : $k^{\text{ième}}$ modalité de basse résolution spatiale.	$B_{kl} = (\text{ex})$ modalité bleue originale à 2.8 m .
$(B_{kl})^{interp}_0$: modalité B_{kl} rééchantillonnée pour s'adapter à la taille de la modalité A_0 .	$(B_{kl})^{interp}_0$: modalité bleue rééchantillonnée pour atteindre la taille (H_1, H_2) de la modalité A_0 .

B_{k0}^* : résultat de la fusion : $k^{\text{ième}}$ modalité synthétisée à la haute résolution.	$B_{k0}^* = B_{1,0}^*$ et $res0 = 0.7$ m
B_l : ensemble des N modalités MS originales.	$B_l = \{B_{1,l}, B_{2,l}, B_{3,l}, B_{4,l}\}$
$(B_l)^{interp}_0$: ensemble des N modalités MS originales rééchantillonnées.	$(B_l)^{interp}_0 = \{(B_{1,l})^{interp}_0, (B_{2,l})^{interp}_0, (B_{3,l})^{interp}_0, (B_{4,l})^{interp}_0\}$
B_0^* : ensemble des N modalités MS fusionnées.	$B_0^* = \{B_{1,0}^*, B_{2,0}^*, B_{3,0}^*, B_{4,0}^*\}$
$B_l(i, j)$: vecteur multispectral formé par les valeurs des N modalités MS originales au pixel (i, j) .	$B_l(i, j) = \{B_{1,l}(i, j), B_{2,l}(i, j), B_{3,l}(i, j), B_{4,l}(i, j)\}$
$(B_l)^{interp}_0(i, j)$: vecteur multispectral formé par les valeurs des N modalités MS originales rééchantillonnées au pixel (i, j) .	$(B_l)^{interp}_0(i, j) = \{(B_{1,l})^{interp}_0(i, j), (B_{2,l})^{interp}_0(i, j), (B_{3,l})^{interp}_0(i, j), (B_{4,l})^{interp}_0(i, j)\}$
$B_0^*(i, j)$: vecteur multispectral formé par les valeurs des N modalités MS fusionnées au pixel (i, j) .	$B_0^*(i, j) = \{B_{1,0}^*(i, j), B_{2,0}^*(i, j), B_{3,0}^*(i, j), B_{4,0}^*(i, j)\}$
(L_1, L_2) : taille des modalités MS originales.	(256, 256)
(H_1, H_2) : taille des modalités MS fusionnées et taille de la modalité A_0 .	(1024, 1024)

Tableau 1.2 : récapitulatif des notations

La figure 1.7 suivante explique graphiquement la définition d'un vecteur multispectral. Un vecteur multispectral est formé par les valeurs des N modalités MS au pixel (i, j) . Dans la littérature, de nombreux termes sont utilisés pour désigner un vecteur multispectral : spectre, signature spectrale, vecteur spectral, N -*uplets* ou encore vecteur multimodal.

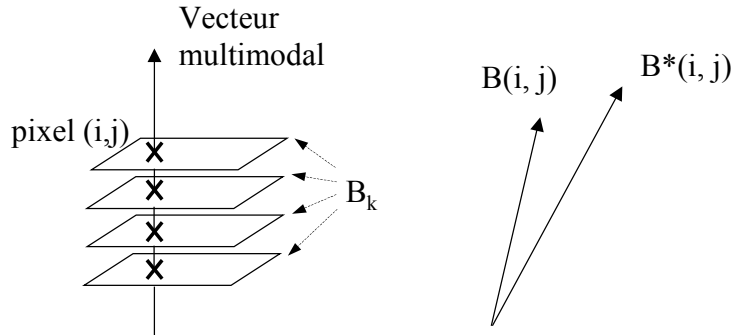


Figure 1.7 : définition du vecteur spectral. On définit par $B(i, j)$ un vecteur MS de référence, et $B^*(i, j)$ le vecteur MS correspondant fusionné.

La propriété fondamentale dit que la fusion doit viser à construire, à partir des représentations originales A et B , des images synthétiques B_{k0}^* qui soient le plus proche possible des images B_{k0} si elles étaient disponibles à cette résolution spatiale $res0$. Les méthodes doivent alors procéder à une transformation de haute qualité du contenu multispectral de B_l lorsque l'on augmente la résolution spatiale de $res1$ vers $res0$ (Wald *et al.* 1997, Wang *et al.* 2005). La fusion d'images revient à estimer la fonction f telle que :

$$B_{k0}^* = f(B_{kl}, A_0) \quad \text{Eq. 1.1}$$

avec B_{k0}^* proche de B_{k0} , ensemble inconnu.

II Qualité

2.1 Introduction

2.1.1 Les enjeux de la qualité des images fusionnées

Les logiciels commerciaux proposent de nombreuses méthodes pour la fusion d'images et il n'est pas évident pour des non-spécialistes de sélectionner une méthode plutôt qu'une autre pour un cas donné. L'utilisateur veut naturellement produire une image de bonne qualité. Cependant, la notion de qualité des produits de fusion est difficile à appréhender puisque chaque acteur possède sa propre vision de la définition d'une image fusionnée de bonne qualité.

Par exemple, un utilisateur qui souhaite procéder à un post-traitement de type classification ou segmentation, n'aura pas les mêmes critères d'évaluation du produit fusionné qu'un autre utilisateur appliquant un algorithme de détection de cible. Les utilisateurs, ou toute personne ayant recours à une méthode de fusion tels que les producteurs d'images fusionnées, formulent des requêtes pour l'image qu'ils doivent fusionner. Une requête peut par exemple favoriser une bonne restitution des couleurs en limitant la distorsion radiométrique plutôt qu'un bon rendu visuel au niveau des hautes fréquences spatiales de l'image fusionnée, ou inversement. Dans ce cas, l'utilisateur souhaite alors connaître *a priori* les avantages et les défauts des différentes méthodes de fusion mises à disposition par les fournisseurs de logiciel de traitement d'images satellite (comme Erdas ou ENVI). Pour cela, il faudrait que ces méthodes aient préalablement été testées sur un nombre important d'images différentes, possédant des ratios et des résolutions spatiales différents, et ce avec plusieurs distances quantifiant la qualité, afin que les utilisateurs puissent se construire une première idée de la qualité des produits délivrés.

La connaissance des points forts et points faibles d'une méthode ne peut être fiable que si un cadre formel d'évaluation de la qualité est établi et respecté. De la même manière, ce cadre est nécessaire une fois le produit de fusion synthétisé. Les utilisateurs doivent avoir les moyens de juger la qualité effectivement obtenue par une liste limitée d'indices pertinents. Si cette liste est trop longue, ils risquent de s'égarer dans leur démarche d'estimation de la qualité. Cette qualité, dite *a posteriori*, est celle à laquelle doivent aussi faire face les développeurs de méthode de fusion. L'innovation implique évidemment qu'il n'existe aucune connaissance *a priori*, et il va alors falloir la positionner par rapport aux méthodes de fusion existantes. Par conséquent, il est nécessaire d'avoir recours à ce même cadre normatif. L'objectif est à terme de venir enrichir la base de données des méthodes de fusion et de pouvoir orienter les futurs utilisateurs de la méthode vers les résultats auxquels ils peuvent s'attendre pour une catégorie donnée d'images.

Pour le moment, seuls quelques auteurs ont entrepris la démarche d'établir un protocole normalisé d'évaluation de la qualité. La lecture de la bibliographie nous enseigne qu'il est judicieux de chercher à harmoniser la manière d'estimer la qualité des images fusionnées afin d'éviter le recours à certaines démarches maladroites, pourtant rentrées dans l'usage courant. Par exemple, un développeur de méthodes s'attache à valider la méthode de fusion qu'il vient de développer. Le nombre d'images testées importe peu puisque la finalité n'est pas l'image, mais la méthode utilisée pour l'obtenir. Seules sont conservées les images et les distances qui sont pertinentes vis-à-vis de la méthode et permettent de la valider. Chaque individu pioche dans la diversité des outils à sa disposition et choisit les démarches ou indices qui le servent le mieux.

2.1.2 Les paramètres qui influencent la qualité des produits fusionnés

Tout d'abord, le résultat de la fusion est sensible à la variété des thèmes présents dans la scène acquise par le capteur satellite. Ce premier point nous permet d'introduire les notions d'objets résolus/non-résolus et de résolution caractéristique. Par exemple, les vagues à la surface de l'océan ou les haies rectilignes des exploitations agricoles ne sont pas résolues à des résolutions de l'ordre de la dizaine ou de la centaine de mètres. Ces structures ne sont résolues que sur une plage étroite en fréquence spatiale. La résolution caractéristique d'une région correspond à une bande de fréquence spatiale qui dépend de la taille des objets présents. Ainsi, un procédé de fusion appliqué à une zone homogène d'un point de vue radiométrique telle que de larges exploitations agricoles, des régions désertiques ou un océan, a de fortes chances d'être très performant. Par contre, à résolution spatiale identique, fusionner les zones urbaines ou les petites parcelles agricoles font partie des cas les plus difficiles. De telles scènes exhibent un grand nombre d'objets de tailles très différentes qui s'intersectent : du large entrepôt au petit pavillon résidentiel, et du parc communal à l'allée régulière de peupliers longeant certaines routes. Par conséquent, comme le spectre de ces paysages est beaucoup plus complexe, le processus de fusion a une probabilité plus importante de donner des résultats plus médiocres.

La qualité du produit fusionné varie aussi en fonction de la résolution spatiale des images MS à fusionner. Si on choisit l'exemple d'une zone urbaine imagée à des résolutions de 80 et 40 m, les deux images ne montreront pas de différences flagrantes. Par exemple, Wald et Ranchin (1995) ont trouvé que la perte d'informations, exprimée grâce à la variance, pour une image montrant la ville de Barcelone (Espagne) en passant de la résolution 40 à 80 m était d'approximativement 20 %. Cette perte est faible et on peut s'attendre alors à ce que le procédé de fusion donne de bons résultats pour une fusion de 80 à 40 m. A une meilleure résolution spatiale, la situation peut s'avérer complètement différente. En effet, pour de l'imagerie de l'ordre du mètre, comme par exemple pour les satellites Ikonos et Quickbird, un grand nombre de structures peut disparaître en passant d'une résolution spatiale vers une résolution plus grossière. L'algorithme de fusion peut rencontrer plus de difficultés pour atteindre la même efficacité.

L'incidence de ces facteurs sur la qualité résultante limite, de manière évidente, la validité du bilan « a priori » aussi complètes que soient les informations à propos de la méthode employée pour une catégorie d'images. La variabilité des images est infinie et une méthode de fusion ne peut être testée sur l'intégralité des images possibles, voire dans d'autres types d'imagerie, telles que l'imagerie médicale ou météorologique.

Notre objectif est de définir un protocole d'évaluation de la qualité des produits de fusion qui respecte la propriété fondamentale des produits de fusion. Nous avons voulu un protocole objectif, générique et indépendant de l'algorithme de fusion ou du type d'image à synthétiser.

2.2 Proposition d'un nouveau protocole d'évaluation de la qualité des produits de fusion

2.2.1 De la littérature aux propriétés

Définir un protocole consiste à définir un ensemble de règles et de procédures, suffisamment globales pour prendre en compte un maximum de cas possibles. Un tel cadre est complexe à définir et a fait l'objet de plusieurs travaux et tentatives de définition. Le protocole que nous proposons se base sur l'étude critique des travaux de ses prédécesseurs pour préserver les éléments essentiels et s'affranchir de leurs limitations.

La bibliographie nous permet de lister et d'identifier les aspects essentiels à vérifier. Les auteurs ont souvent tendance à privilégier certaines caractéristiques des produits de fusion par rapport à d'autres dans leur démarche d'évaluation de la qualité. Certains cherchent à quantifier la qualité

géométrique des images, comme le recours à la *PSF* (*Point Spread Function*, Li 2000) ou à un filtrage passe-haut du produit de fusion (Zhou 1998). IRARS (1996) a privilégié l'identification visuelle des différentes régions présentes dans le produit de fusion vis-à-vis d'une grille de critères aussi bien géométriques que radiométriques. La distorsion radiométrique est habituellement mise en évidence grâce à l'utilisation de distances telles que le biais ou la racine carrée de la moyenne quadratique des différences, notée *RMSE* pour l'acronyme en anglais de *Root Mean Square Error* (Wald *et al.* 1997, Alparone *et al.* 1998). Les propriétés multispectrales des images sont aussi considérées dans un grand nombre de travaux, soit sous la forme d'une analyse visuelle de compositions colorées (IRARS 1995, Vrabel 2000, Boissezon et Laporterie 2002), soit grâce à l'intervention de critères quantitatifs multimodaux (Wald *et al.* 1997, Aiazzi *et al.* 2003). Notre protocole doit pouvoir couvrir l'ensemble des notions citées précédemment et en proposer une organisation structurée sous forme de propriétés à vérifier. Bien qu'incomplètes, Wald *et al.* (1997) et Li (2000) ont constitué deux références de choix pour notre étude en posant d'excellentes bases pour la structure d'un bon protocole.

Nous avons défini deux propriétés principales appelées « propriété de cohérence » et « propriété de synthèse ».

- ◆ La propriété de cohérence doit permettre de protéger l'ensemble initial lors de l'augmentation de la résolution spatiale des images MS. En d'autres termes, si une méthode de fusion est appliquée, il faut vérifier que par dégradation spatiale, nous sommes capable de retrouver l'ensemble original. Cette propriété du produit de fusion pourrait apparaître comme triviale. Or de nombreuses méthodes très répandues comme *IHS* ne vérifient pas cet aspect et ne proposent qu'une très faible cohérence avec l'ensemble des images originales.
- ◆ La propriété de synthèse correspond à la retranscription mathématique de notre point de départ : elle traduit le respect de la propriété fondamentale des produits de fusion : le produit synthétisé doit ressembler à celui qui aurait pu être acquis par le capteur MS correspondant si celui-ci possédait une meilleure résolution spatiale.

Les protocoles existants abordent l'évaluation des images fusionnées de deux manières différentes : monomodale (modalité par modalité) et multimodale (toutes modalités confondues, pouvant aussi être appelé 'multispectrale'). Pour l'aspect monomodal, chaque modalité fusionnée est comparée à sa référence correspondante ; le bilan est établi en ne tenant pas compte des autres modalités. Pour l'aspect multimodal, on vérifie si les propriétés multimodales de l'ensemble fusionné sont proches de celles de référence. Par exemple, il est possible d'utiliser des distances qui donneraient une idée globale de la qualité pour l'ensemble des modalités telles que la somme des *RMSE* (Munechika *et al.* 1993) ou l'Erreur Relative Globale Adimensionnelle de Synthèse (*ERGAS*, Wald 2000, 2002). Les deux propriétés de cohérence et de synthèse doivent être considérées selon ces deux points de vue.

2.2.2 La propriété de cohérence

Toute image synthétique B_{k0}^* , une fois dégradée à sa résolution spatiale originale I , doit être aussi proche que possible de l'image originale B_{kl} . Cette propriété doit être vérifiée pour les modalités prises individuellement ainsi que pour toutes les modalités confondues.

Pour l'aspect monomodal, la propriété peut se traduire par le critère suivant :

$$D_I(B_{kl}, (B_{k0}^*)_I) < \varepsilon_{Ik} \quad \text{Eq. 2.1}$$

où $(B_{k0}^*)_I$ représente l'image B_{k0}^* rééchantillonnée à la résolution I , et D_I est une distance entre B_{kl} et $(B_{k0}^*)_I$. Une distance possible est le biais entre ces deux images, relatif à la moyenne de l'image de référence. Le seuil associé à cette distance peut être fixé à une valeur de 0.5 %. Si une modalité présente un biais inférieur à 0.5 %, le produit de fusion sera considéré comme étant de bonne qualité. D'autres distances peuvent aussi être utilisées telles que la différence en variance relativement à la variance de la référence (variance de B_{kl}), ou encore l'entropie, qui correspond à la quantité d'information présente dans une image (Wald *et al.* 1997). L'approximation introduite par le

rééchantillonnage de B_{k0}^* à $(B_{k0}^*)_I$ doit être prise en compte : la limite ε_{Ik} est déterminée par le degré de précision requis. ε_{Ik} doit être faible pour chaque modalité.

Cette propriété s'applique aux modalités prises individuellement. Elle n'implique pas forcément une synthèse précise des propriétés multispectrales du jeu B en augmentant la résolution spatiale. C'est pourquoi l'aspect multimodal doit aussi être vérifié par un critère comme le suivant :

$$D_2(B_I, (B_0^*)_I) < \varepsilon_2 \quad \text{Eq. 2.2}$$

où D_2 représente une distance entre B_I et $(B_0^*)_I$. Une distance multimodale peut être l'*ERGAS* avec un seuil de qualité associé ε_2 de 3. Si un produit de fusion présente un *ERGAS* inférieur à 3, il est alors de bonne qualité (Wald 2002). Comme précédemment, l'approximation introduite par le rééchantillonnage de B_0^* vers $(B_0^*)_I$ doit être prise en compte : ε_2 doit être faible pour toutes les modalités pour assurer la similarité entre les deux ensembles d'images.

L'aspect monomodal de la propriété de cohérence est le même que la première propriété définie antérieurement par Wald *et al.* (1997) et par Li (2000). Notre propriété de cohérence étend ces travaux.

2.2.3 La propriété de synthèse

Toute image synthétique B_0^* doit être aussi proche que possible de celle qu'observerait le capteur correspondant s'il possédait la résolution spatiale 0. Cette propriété doit être vérifiée pour les modalités prises individuellement ainsi que pour toutes les modalités confondues.

Pour l'aspect monomodal :

$$D_3(B_{k0}, B_{k0}^*) < \varepsilon_{3k} \quad \text{Eq. 2.3}$$

où D_3 est une distance entre B_{k0} et B_{k0}^* pour la modalité k . L'ensemble des distances D_3 peut tout à fait être identique à celui de D_I , ou différent selon les objectifs de l'application.

Pour l'aspect multimodal :

$$D_4(B_0, B_0^*) < \varepsilon_4 \quad \text{Eq. 2.4}$$

où D_4 est une distance entre B_0 et B_0^* . L'ensemble des distances D_4 peut être soit le même, soit différent de l'ensemble des distances D_2 selon l'application choisie.

Cette propriété de synthèse est identique aux deuxième et troisième propriétés proposées par Wald *et al.* (1997), et pour l'aspect multimodal à la troisième propriété de Li (2000). Elle unifie les travaux des premiers et étend ceux du second. La seconde propriété de Li (2000) traite d'un aspect particulier des produits fusionnés. Cette caractéristique monomodale est implicitement contenue dans l'équation 2.3.

Etablir un bilan de qualité implique la comparaison entre le produit fusionné (ses propriétés ou des quantités dérivées) et une référence. Cette évaluation s'effectue de manière qualitative (ou analyse visuelle) et de manière quantitative. Une bonne qualité sera obtenue si le produit est proche de sa référence. Concernant la propriété de cohérence, l'ensemble de référence est immédiat et correspond à l'ensemble MS initial B_I . Par contre, aucune référence ne s'impose à nous de manière évidente pour la propriété de synthèse ; il faut alors trouver une alternative.

2.3 Le problème de la référence

Le problème de la sélection de la référence n'est pas nouveau. Il a déjà été souligné par plusieurs auteurs. Cependant, l'étude de la bibliographie montre un certain nombre de lacunes dans le choix de cette alternative.

2.3.1 Lacunes rencontrées dans la littérature

2.3.1.1 Images MS originales

Parmi les approches les plus rencontrées dans la littérature, une d'entre elles consiste à utiliser l'image de basse résolution B_{kl} rééchantillonnée à la haute résolution $res0$ ($(B_{kl})^{interp_0}$). Cette démarche a été suivie dans les travaux suivants : Chavez *et al.* (1991), Yocky (1996), Alparone *et al.* (1998), Terretaz (1998), Liu (2000), Zhijun *et al.* (2004), et cette liste n'est pas exhaustive. Il existe un cas où les images interpolées peuvent être des références ; si les objets de la scène sont spatialement homogènes pour des échelles comprises entre $res0$ et $res1$, n'importe quelle méthode de fusion sensée fournira de bons résultats à $res0$. Mais dans ce cas, l'intérêt de la fusion d'images est remis en cause puisqu'un simple rééchantillonnage délivre aussi des résultats satisfaisants. Nous excluons ce cas particulier.

Dans toutes les autres situations, les images suréchantillonnées ne sont pas une alternative satisfaisante. Il existe de nombreuses façons de suréchantillonner une image : par exemple, le logiciel ENVI propose les plus proches voisins, un suréchantillonnage bilinéaire ou encore bicubique. Ce sont des approximations du suréchantillonnage idéal qui correspond à une fonction porte dans le domaine de Fourier. Cet opérateur se traduit par la fonction sinus cardinal ($\sin(x)/x$). Or, ce filtre est de support infini, ce qui entre en contradiction avec l'image qui est de support fini. Chaque opérateur donne une image différente, ce qui implique que pour une même image, des utilisateurs pourraient avoir des références différentes, ce qui est faux puisque par définition, une référence est unique.

En outre, les images rééchantillonnées ne sont pas représentatives de ce que pourraient observer les capteurs correspondants s'ils possédaient la haute résolution $res0$ puisqu'à aucun moment les caractéristiques du capteur ne sont prises en compte. Ces images ne peuvent en aucune sorte constituer des références valides. Cette approche n'est donc pas correcte et ne devrait pas être utilisée. Si tel était le cas, il y aurait un paradoxe : si des images rééchantillonnées sont supposées être les références, quel est l'intérêt de la fusion d'images ?

2.3.1.2 Quantités statistiques des images originales

D'autres protocoles ont essayé de contourner ce problème de référence en utilisant des quantités statistiques appliquées aux données originales et aux données fusionnées. Par exemple, Garguet-Duport *et al.* (1996) comparent les histogrammes des images B_1 à ceux des images B_0^* en affirmant qu'ils doivent être les plus ressemblants possibles. La comparaison d'histogrammes est un bon estimateur de la qualité des images car c'est un outil connu et facile à manipuler. Cependant, l'effet du changement de résolution spatiale sur les histogrammes et sur n'importe quelle quantité statistique de manière générale, ne doit pas être négligé. En effet, le moyennage spatial a tendance à « rogner » les extrémités de l'histogramme, puisque par définition, cette opération tend à regrouper les comptes numériques autour de la moyenne. En conséquence, le produit fusionné devrait posséder une dynamique plus large que celle du signal d'entrée B_1 .

Le tableau 2.1 est une illustration des variations des quantités statistiques avec la résolution spatiale. Nous avons calculé la moyenne et la variance d'une image Pan d'une zone urbaine, acquise par le satellite Quickbird tout d'abord à sa résolution d'acquisition (0.7 m) puis dégradée huit fois (5.6 m). On remarque que si la moyenne est conservée, la variance est pratiquement divisée par un facteur 2 en réduisant la résolution. Cette observation est similaire à celle apportée à propos de l'image de Barcelone dans le second paragraphe de la partie 2.1.2. Il existe donc une différence notable entre les quantités statistiques de B_1 et celles des images B_0^* .

	Modalité à $res0 = 0,7$ m	Même modalité dégradée à $res1 = 5,6$ m
Moyenne	249	249
Variance	$284 \cdot 10^2$	$152 \cdot 10^2$

Tableau 2.1 : évolution de grandeurs statistiques avec le changement d'échelle (valeurs en niveaux de gris)

Plusieurs travaux ont exploré la variation de grandeurs physiques avec les échelles. Par exemple, Kristjansson (1991) s'est intéressé à l'influence de la résolution dans la prédiction météorologique et les modèles de climat. Raffy (1993) a établi les fondamentaux mathématiques qui ont permis d'expliquer ces variations sur des cas simples. Les résultats obtenus par Raffy sont très semblables à ceux délivrés par Welch *et al.* (1989, fig. 3), Lillesand et Kiefer (1994, fig. 7.53), Woodcock et Strahler (1987), Kong et Vidal-Madjar (1988). Tous ces travaux ont montré expérimentalement que la qualité est une fonction complexe et aléatoire de la résolution (Wald *et al.* 1997). Cela signifie que l'on ne doit pas essayer de comparer des propriétés statistiques du produit de fusion et de l'image originale : histogrammes, fréquences cumulées, variances, entropies, coefficients de corrélation... Donc les protocoles d'évaluation de la qualité des produits de fusion basés sur la comparaison de quantités statistiques ne devraient pas être utilisés.

2.3.1.3 L'image de haute résolution spatiale comme référence

Chavez (1991), Zhou (1998), Zhang (1999), Li (2000), Zhijun *et al.* (2004), Vijayaraj *et al.* (2004) et Wang *et al.* (2005) ont fait la démarche de choisir comme référence, non plus les modalités MS originales, mais l'image Pan par l'intermédiaire de différents opérateurs (filtrage passe-haut, coefficient de corrélation, profil d'une ligne image...). Cette fois, les deux images ont bien la même résolution spatiale et l'objectif est de venir comparer les hautes fréquences présentes dans les deux images.

Cependant, la similarité attendue implique qu'aucun changement des hautes fréquences n'intervient lorsque l'on passe d'une modalité à une autre. Or nous avons vu dans le chapitre d'introduction des exemples où, localement des hautes fréquences apparaissent dans certaines modalités et sont absentes des autres à cause de leur bande spectrale d'acquisition. D'ailleurs, certains auteurs reconnaissent être gênés par cette démarche, sans pour autant en étudier la contradiction. Par exemple, Li (2000) et Vijayaraj *et al.* (2004) précisent que le coefficient de corrélation entre les hautes fréquences du Pan et de B_0^* doit être élevé sans l'être « trop ». La modalité Pan ne peut pas constituer de référence pertinente pour les modalités MS.

La référence doit prendre en compte certaines informations provenant de la modalité de haute résolution spatiale concernant les hautes fréquences spatiales qui vont être introduites, mais doit aussi prendre en compte une partie d'innovation qui est propre à la modalité traitée. Le besoin de référence à haute résolution spatiale est une contrainte stricte que nous avons cherché à alléger. Voilà pourquoi nous avons proposé un protocole qui a la particularité de toujours se ramener aux modalités MS d'origine, puisque ce sont les seules références valables.

2.3.2 Proposition de référence et vérification de la propriété de synthèse

Considérant que seules les images originales B_{kl} constituent des références vraisemblables, nous proposons un changement d'échelle, et de travailler à plus basse résolution, comme l'ont proposé Munechika *et al.* (1993), Mangolini (1994), Wald *et al.* (1997), Aiazzi *et al.* (2002), De Boissezon et Laporterie (2002), Buntikov et Bretschneider (2004), Vijayaraj *et al.* (2004). Les deux ensembles d'entrées B_2 et A_1 sont créés par rééchantillonnage des ensembles d'origine B_1 et A_0 avec la très basse résolution $res2$ définie par $res2/res1 = res1/res0$. La méthode de fusion est appliquée aux deux ensembles B_2 et A_1 . On obtient alors des images synthétisées B_1^* à la résolution spatiale des images MS d'origine, qui servent donc de référence. Un bilan de la qualité obtenue à cette résolution spatiale $res1$ peut être établi. Finalement, on suppose que la qualité obtenue pour le produit de fusion B_1^* est

proche de celle que l'on obtiendrait pour le produit B_0^* si on possédait la référence à $res0$. Il s'agit d'une hypothèse que nous appelons « hypothèse d'extrapolation » et qui est discutée ultérieurement.

Notons que, excepté pour Mangolini (1994) et Wald *et al.* (1997), le protocole d'évaluation de la qualité des produits fusionnés se limitait à la vérification de la propriété de synthèse, et ne vérifiait donc pas la cohérence des résultats avec l'ensemble original.

2.3.3 Application du protocole

La formulation des propriétés de cohérence et de synthèse combinée à la proposition de référence dans le cas de la propriété de synthèse permet de définir les étapes d'évaluation suivantes.

Etape 1 : la méthode de fusion est appliquée sur les deux ensembles d'images initiaux : A_0 et B_I . Il en résulte ainsi un nouveau jeu d'images B_0^* . La première étape, qui fait partie de la vérification de la propriété de synthèse, consiste alors à visualiser les produits fusionnés afin de déceler les artefacts flagrants. Cette visualisation s'effectue en comparaison avec les ensembles originaux A_0 et B_{kl} pour chaque modalité k , puis en couleurs en utilisant une ou plusieurs compositions colorées. Notons que les résultats de cette visualisation sont incomplets puisque la référence n'existe pas.

Etape 2 : vérification de la propriété de cohérence : l'ensemble B_0^* est dégradé jusqu'à la résolution I (B_{k0}^*)_I. On analyse visuellement l'écart entre (B_{k0}^*)_I et B_{kl} pour chaque modalité k , les images sont en niveaux de gris. De la même façon, on compare les images couleurs (B_0^*)_I et B_I . Enfin, on mesure la qualité monomodale et multimodale grâce à des distances adaptées. Les résultats obtenus pour ces distances sont comparés aux critères définis pour chaque distance.

Etape 3 : la vérification de la propriété de synthèse se poursuit avec l'utilisation des distances. Pour cela, un changement d'échelle est opéré. La méthode de fusion est alors appliquée aux images A_I et B_2 qui proviennent de la dégradation des images originales A_0 et B_I . Nous obtenons alors des images B_I^* que l'on compare aux images originales B_I encore une fois grâce à une évaluation visuelle de la qualité, suivie par le calcul de distances monomodales visant à comparer B_{kl}^* et B_{kl} , puis multimodales avec la comparaison de B_I^* avec B_I . Les résultats obtenus pour ces distances sont comparés aux critères définis pour chaque distance.

Etape 4 : on admet alors que le bilan de qualité établi pour les produits de fusion B_{kl}^* et l'ensemble global B_I^* est proche de celui que l'on aurait pu établir à la résolution supérieure $res0$ si les images B_{k0} , et donc l'ensemble B_0 , avaient été disponibles.

Si la référence existe à $res0$, alors les étapes se simplifient. La première étape du protocole qui constitue la visualisation à haute résolution des produits peut alors s'effectuer en parallèle avec les modalités MS de référence à haute résolution, sur lesquelles les distances pour la vérification de la propriété de synthèse peuvent directement être appliquées. Par conséquent, les étapes 3 et 4 de ce protocole disparaissent.

2.3.4 Conclusion et limites du protocole

Dans le cadre de la synthèse d'images que nous nous sommes fixé, le protocole proposé est générique car il est capable de traiter n'importe quel paysage, indépendamment du ratio ou de la méthode de fusion utilisée. Notre objectif est de réunir les acteurs de la fusion d'images autour d'un protocole le plus général possible. Pour le moment, ce protocole a été présenté lors du colloque EARSeL 2004 (Thomas et Wald 2004) et a reçu des remarques favorables de l'ensemble de la communauté de la fusion présente. Un protocole unanimement reconnu et utilisé permet de rendre comparables les bilans de qualité. De plus, développeurs de méthodes de fusion et utilisateurs utiliseraient un langage unique, ce qui améliorerait la formulation des requêtes par les utilisateurs et leurs satisfactions par les développeurs. La dissémination de l'information serait plus importante, ce qui entraînerait une meilleure valorisation scientifique dans le domaine de la fusion et une meilleure exploitation industrielle des résultats.

2.3.4.1 L'influence du sous-échantillonnage

Cependant, il faut souligner l'influence du choix de l'opérateur de sous-échantillonnage lors de la vérification de la propriété de synthèse et de la propriété de cohérence. Le sous-échantillonnage altère le contenu des images ; Raffy (1994) montre que si l'on applique divers opérateurs (opérateur de dégradation/moyennage) à une image haute résolution, les images basse résolution qui en résultent peuvent s'avérer être totalement différentes. Or, ce processus est vrai à l'inverse, de la basse vers la haute résolution. Cette influence a aussi été soulignée dans les travaux de Ranchin (1993), Garguet-Duport *et al.* (1996), Nuñez *et al.* (1999), Price (1999), Boissezon et Laporterie (2002), Zhijun *et al.* (2004). Que ce soit pour sur- ou sous-échantillonner une image, l'opérateur doit simuler ce qui serait observé par le capteur correspondant s'il possédait une résolution plus faible. Et l'image obtenue est donc unique. Malheureusement une fois encore, un filtre de dégradation idéal n'existe pas puisque les images sont de support fini. L'influence du sous-échantillonnage est importante sur la qualité obtenue. Il s'agit d'un des points faibles de notre protocole.

Cependant, il est possible d'utiliser des filtres qui permettent de limiter les effets néfastes de la dégradation. Ces filtres doivent avoir la caractéristique de séparer correctement les échelles, c'est-à-dire de séparer les structures de petites tailles de celles de tailles plus importantes. Ranchin (communication personnelle) a comparé plusieurs opérateurs appliqués à diverses scènes tels que le sinus cardinal tronqué par la fonction d'apodisation de Hanning de taille 13x13, la fonction de Shannon tronquée, une Spline bi-cubique, et l'ondelette de Daubechies. Les images résultantes montrent des écarts relatifs de l'ordre de quelques pourcents. La conclusion qui a découlé de cette étude est que l'influence du filtrage existe, mais peut toutefois être considérée comme négligeable si le filtre est suffisamment approprié. Une Spline bi-cubique offre souvent un bon compromis.

Les travaux de Aiazzi *et al.* (2002) ont montré que des opérateurs qui simulent la Fonction de Transfert de Modulation (*FTM*) du capteur donnent de très bons résultats. Ce point est très important puisqu'il caractérise la qualité intrinsèque du capteur qui a acquis l'image. Il fera l'objet du chapitre 4.

2.3.4.2 L'extrapolation du bilan de qualité

Le second point faible du protocole, dès lors que nous ne disposons pas d'image de référence à *res0*, est l'hypothèse d'extrapolation, à savoir que la qualité obtenue par une fusion à une échelle plus basse est proche de celle obtenue à haute résolution. Nous espérons que, tout au moins, le bilan de qualité établi à *res1* est au moins un minorant de celui à *res0*. Mais cette hypothèse ne peut être démontrée théoriquement. Plusieurs auteurs (Hendersen-Sellers *et al.* 1985, Meentemeyer *et al.* 1987, Grégoire *et al.* 1993, Aiazzi *et al.* 2003, Garzelli *et al.* 2004) ont controversé cette idée en soulignant par des exemples, la variabilité et la non-prédictibilité du bilan de qualité avec les échelles. De plus, en montrant que les modalités MS à basse résolution spatiale ne pouvaient servir de référence à celles produites à haute résolution, nous nous sommes appuyés sur plusieurs travaux indiquant expérimentalement que la qualité est une fonction complexe et aléatoire de la résolution.

Cependant, l'expérience montre qu'il existe de nombreux cas pour lesquels l'hypothèse est vérifiée. Nous avons donc mené une étude dans le chapitre 3 pour essayer de cerner le domaine d'applicabilité de l'hypothèse et donc de notre proposition pour une référence. Cette étude empirique vient compléter la proposition de protocole.

2.4 Les outils de l'évaluation de la qualité des produits de fusion

Pour valider les deux propriétés du protocole proposé, l'utilisateur doit procéder à une analyse visuelle et quantitative des produits fusion.

2.4.1 Analyse visuelle

Préconisée par la littérature et les experts, l'analyse visuelle du résultat et des différences est un outil essentiel. Le recours à l'analyse visuelle peut paraître subjectif car le jugement d'une personne sur le résultat esthétique d'une image va dépendre de ses connaissances, de ses attentes et de l'application considérée. Cependant, l'utilisation du système visuel humain en tant qu'outil de jugement de la qualité n'est pas à négliger ; l'utilisateur pourra voir immédiatement la tendance générale des distorsions de l'image en fonction de ses propres attentes.

Il existe une littérature abondante qui propose diverses classifications pour établir la qualité visuelle des images satellite. Le gouvernement des Etats-Unis (IRARS 1995, 1996) a établi et mis à disposition une échelle appelée *Multispectral Imagery Interpretability Rating Scale* (MIIRS) qui permet de classer une image selon le niveau d'information qui peut être extrait de l'image MS, appelé aussi niveau d'interprétabilité (Vrabel 1996). Un rang dans cette classification est attribué à une image composite. Ce rang est unique pour l'ensemble des modalités MS d'un satellite pour une scène donnée. Dans le cas où le nombre de modalités MS est supérieur à trois, l'espace des données peut être réduit, par exemple en utilisant une transformation en composantes principales. Le satellite Landsat Thematic Mapper délivre sept modalités MS simultanément, soit plus de deux cents combinaisons possibles différentes. Pour réduire la quantité d'information en entrée, il faut choisir qualitativement la combinaison dans laquelle les objets sont le mieux discriminés.

2.4.1.1 Recommandations

L'analyse monomodale des images fusionnées consiste en une comparaison de chaque modalité indépendamment des autres avec la modalité de référence correspondante. L'évaluation visuelle multimodale ou en « couleurs » nécessite la sélection de trois modalités parmi les modalités disponibles. Par exemple, si un capteur possède trois modalités MS dans le visible qui couvre approximativement les bandes spectrales du rouge, du vert et du bleu, il sera possible par une combinaison colorée de retrouver une information colorée assez proche de la réalité. C'est ce que l'on appelle une image en vraies couleurs ou en couleurs naturelles. Une image dite en fausses couleurs peut aussi être visualisée en sélectionnant par exemple la modalité PIR à la place du canal rouge. L'avantage de ce canal sensible à la chlorophylle est de faire ressortir en rouge la végétation d'une scène (le rouge est une couleur saturante pour l'œil humain donc très visible).

Nous souhaitons attirer l'attention sur les outils de visualisation d'images mis à disposition des utilisateurs pour visualiser de manière monomodale ou multimodale. Lors de l'affichage à l'écran d'une image par celui-ci, une table de couleurs ou de niveaux de gris est générée automatiquement pour s'adapter au mieux à la vision humaine, c'est-à-dire en proposant un contraste suffisant permettant la distinction des détails les plus fins de l'image. Que ce soit pour les images couleurs ou les images en niveaux de gris, cette opération s'effectue généralement de manière opaque pour l'utilisateur du logiciel et peut entraîner des erreurs de comparaison. Pour illustrer cette remarque, nous avons choisi une image acquise par le satellite Ikonos au-dessus de parcelles agricoles en périphérie de Hasselt, une ville de Belgique. La figure 2.1a correspond à la modalité Pan d'origine, et celle placée en b est la même image, mais où tous les comptes numériques ont été multipliés par 2. Le logiciel de visualisation est ENVI. En apparence, ces deux images sont identiques. La figure 2.1c quant à elle correspond à une image que nous appellerons « mosaïque ». Cette solution a été proposée par Albuison (1993). Elle résulte de la création d'une image supplémentaire comprenant sur la gauche la référence, et sur sa droite, la modalité synthétisée. Une table de couleurs (ici niveaux de gris) unique a été affectée à cette image. La différence entre les deux images apparaît cette fois-ci très nettement.

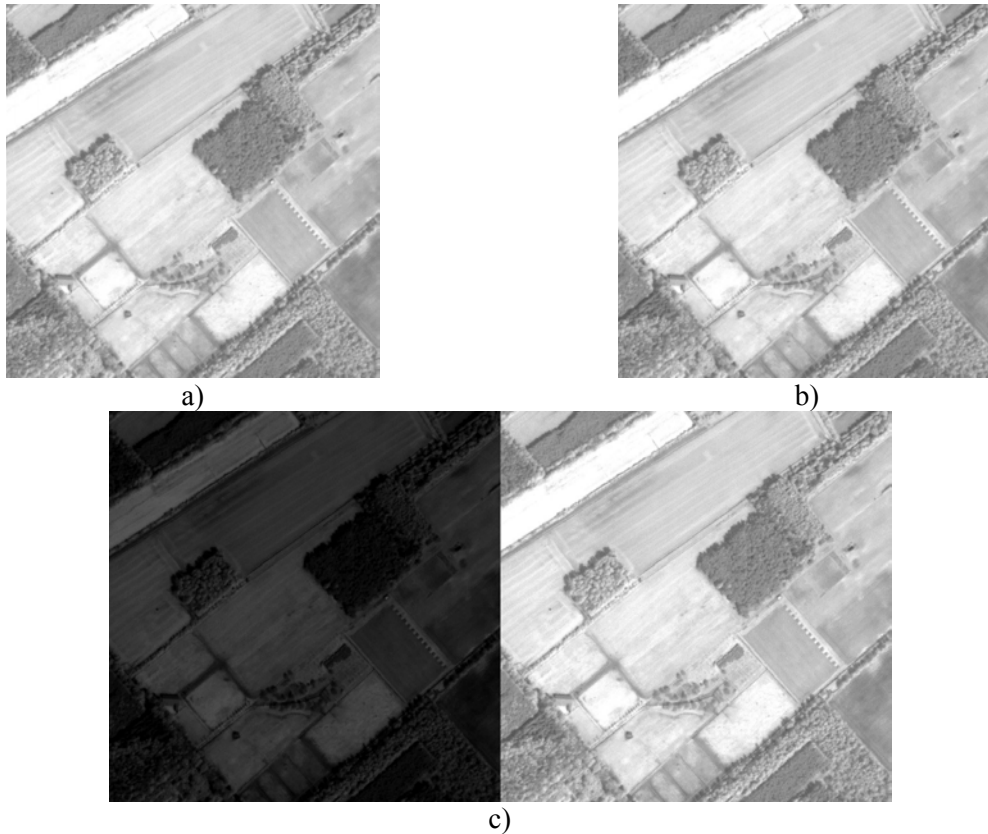


Figure 2.1 : méthode de visualisation de deux images : création d'une image intermédiaire (ou mosaïque) qui contient les deux images à comparer.

En dépit des traitements inhérents au logiciel de visualisation, d'autres éléments peuvent influencer la visualisation d'une image. La limitation du nombre de couleurs perceptible sur un même écran par l'œil humain est une première source d'erreur. Une deuxième source d'erreur est la différence entre le nombre de couleurs de l'image MS couleurs et celui qu'une carte graphique est capable de traiter. Si le nombre maximum de couleurs traitées par la carte graphique est plus faible que celui de l'image, la carte graphique opère une compression de l'information contenue dans l'image avant d'attribuer à un pixel une couleur. Cette compression automatique se fait par analyse de quantités statistiques.

En conclusion, une analyse visuelle efficace est facilitée si le logiciel attribue la même table de couleurs aux deux images à comparer. Par conséquent, nous recommandons l'utilisation de mosaïque pour l'analyse mono- et multimodale.

2.4.1.2 *Evaluation visuelle des deux propriétés*

Comme nous l'avons expliqué dans les étapes de l'application du protocole, l'évaluation visuelle de la qualité doit s'effectuer de la même manière que pour les distances. C'est à dire:

- ♦ $(B_{k0}^*)_I$ et B_{kl} pour la propriété de cohérence monomodale,
- ♦ $(B_0^*)_I$ et B_I pour la propriété de cohérence multimodale,
- ♦ B_{k0}^* , B_{kl} et A pour la propriété de synthèse monomodale,
- ♦ B_0^* , B_I et A pour la propriété de synthèse multimodale.
- ♦ B_{kl}^* et B_{kl} pour la propriété de synthèse monomodale,
- ♦ B_I^* et B_I pour la propriété de synthèse multimodale.

L'analyse visuelle permet de localiser et d'identifier des défauts. Ces défauts peuvent affecter la qualité image (trop de piqué, aspect pixellisé ...), ou la géométrie (déformation d'éléments linéaires, bavé des contours, agrégation d'objets ...) ou encore la radiométrie (saturation des couleurs, couleur dominante, dynamique ...).

Nous attirons l'attention sur le fait que nous recommandons de procéder à une analyse visuelle des produits de fusion à haute résolution spatiale. Elle correspond à la première étape de l'application du protocole. Même sans aucune référence directe, l'observation du résultat en parallèle avec l'image Pan et les images MS originales est conseillée car elle permet d'identifier les défauts flagrants. Il s'agit de la manière de procéder pour la plupart des campagnes d'expérimentation (De Boissezon et Laporterie 2002). Nous avons nous-mêmes procédé à deux campagnes d'expérimentation lors de nos travaux de recherche. Ces campagnes consistent à réunir un groupe d'interprètes d'images extérieurs à notre laboratoire pour juger de la qualité des produits de fusion à haute résolution spatiale. Elles nous ont permis d'identifier ou de confirmer les défauts et artefacts de divers produits de fusion obtenus par des algorithmes de fusion existants ou en cours de développement. Néanmoins, les analyses procurées par leur analyse visuelle doivent être manipulées avec une grande attention car :

- ◆ les expérimentateurs n'ont pas procédé à l'analyse visuelle pour la vérification de la propriété de cohérence,
- ◆ ni à celle pour la propriété de synthèse à basse résolution,
- ◆ enfin, ce sont des personnes habituées à un tel exercice et dans leur manière de procéder, ils n'ont pas recours à la création d'une mosaïque. Leurs remarques et annotations sont donc dépendantes des tables de couleurs allouées par l'outil de visualisation (ENVI).

2.4.2 Analyse quantitative

Le protocole doit s'accompagner d'un certain nombre de distances significatives et critères associés. Il est vraisemblable que ces distances sont dépendantes de l'objectif final des utilisateurs (Ranchin et Wald 2003). Cependant, on pourrait s'attendre à ce que plusieurs distances et critères soient suffisamment significatifs pour être adoptés en tant que standards tels que, par exemple, le biais ou la différence entre les variances des deux images à comparer.

Toutes les distances ont leurs qualités et leurs défauts. Exceptées pour quelques exceptions, si les distances sont appliquées indépendamment les unes des autres, elles sont insuffisantes pour décrire la qualité d'un produit fusion dans son intégralité. Voilà pourquoi il est nécessaire de combiner ces distances pour pouvoir obtenir une idée plus complète de la qualité. Pour cela, nous avons parcouru 36 articles et communications et avons recensé trente-neuf distances différentes. Généralement, les auteurs piochent entre deux et quatre distances dans cette diversité sans pour autant s'attarder à savoir s'ils se sont bien attachés à vérifier tous les aspects de la qualité. Plusieurs illustrations simples ainsi que quelques démonstrations mathématiques et études empiriques nous permettent de montrer que certaines distances sont redondantes tandis que d'autres se complètent. Nous proposons une classification des distances en sept catégories. Enfin, nous proposons diverses combinaisons de distances, formant ce que nous appellerons des bilans de qualité.

Nous avons implanté la majorité de ces distances dans un atelier logiciel programmé en IDL (Thomas *et al.* 2005). L'interface graphique a été placée en annexe (annexe 1).

2.4.2.1 Catégorie 1 : « différence entre mesures globales »

Les distances concernent des différences entre des mesures globales appliquées à une modalité synthétisée et son homologue de référence. Ces mesures globales peuvent être par exemple la moyenne ($M(B_k)$), la variance ($V(B_k)$), l'écart type ($\sigma(B_k)$) ou encore l'entropie ($H(B_k)$) (Garguet-Duport *et al.* 1996, Terretaz 1997, Blanc *et al.* 1998, Li 2000, Wald 2002, Ballester *et al.* 2003, Ranchin *et al.* 2003, Rokhmatuloh *et al.* 2003, Alonso Reyes *et al.* 2005a, Acerbi-Junior *et al.* 2006, Nencini *et al.* 2006, Yang *et al.* 2006). Comme cette dernière est rencontrée moins fréquemment dans la littérature, nous précisons son écriture :

$$H = - \sum_{i=\min}^{\max} p(i) \ln(p(i)) \quad \text{Eq. 2.5}$$

où $p(i)$ est la probabilité d'avoir le niveau de gris i . Cette dernière se calcule en divisant le nombre de pixels qui possèdent la valeur i par le nombre total de pixels de l'image. Ainsi $p(i)$ qui est un rapport de nombres de pixel est bien adimensionné.

La distance la plus couramment utilisée est le biais, qui correspond à la différence des moyennes ($(\text{biais}(B_k, B_k^*)) = M(B_k) - M(B_k^*)$). La valeur de cette distance est souvent donnée en valeur relative par rapport à la moyenne de la référence. De même, les différences de variances, d'écart types ou encore d'entropies sont divisées respectivement par la variance ($\text{diffVarRel}(B_k, B_k^*)$), l'écart-type ($\text{diff}\sigma\text{Rel}(B_k, B_k^*)$) et l'entropie ($\text{diffHRel}(B_k, B_k^*)$) de la modalité de référence. Ces distances ont originellement été proposées par Munechika *et al.* (1993) et Wald *et al.* (1997) et sont très souvent utilisées. Une autre distance est la différence entre les modes des histogrammes de chaque image (Ballester *et al.* 2003).

Wald *et al.* (1997) et Wald (2002) ont aussi proposé une distance multimodale qui consiste à calculer la différence du nombre de signatures spectrales entre l'image fusionnée et de référence. Une valeur positive signifie que la fusion n'a pas synthétisée assez de *N-uplets*, alors qu'une différence négative implique une trop grande innovation.

Comme ces distances reposent sur des différences entre grandeurs appliquées aux images de référence et fusionnées, la valeur idéale de chacune de ces distances est 0.

2.4.2.2 Catégorie 2 : « distances appliquées sur l'image de différence »

Les distances de cette catégorie s'appliquent sur l'image de différence entre B_k^* et B_k calculée en chaque pixel (i, j) . L'écart-type relatif à la moyenne de la référence ($\sigma\text{Rel}(B_k, B_k^*)$), (Terretaz 1997, Blanc *et al.* 1998, Li 2000, Wald 2002, Ranchin *et al.* 2003, Rokhmatuloh *et al.* 2003, Alonso Reyes *et al.* 2005a, Thomas and Wald 2006b) et la *RMSE* sont les distances les plus connues. La *RMSE* (Munechika *et al.* 1993, Wald *et al.* 1997, Alparone *et al.* 1998, Bunttilov et Bretschneider 2004, Chen et Varshney 2005, Zheng *et al.* 2005) s'écrit :

$$\text{RMSE}(B_k, B_k^*) = 1/nb\text{Pix} \sqrt{\sum (B_k^*(i,j) - B_k(i,j))^2} \quad \text{Eq. 2.6}$$

Des distances équivalentes au *RMSE* et à l'écart-type sont la variance et l'erreur moyenne quadratique. Ces distances sont généralement divisées par la moyenne de la référence. D'autres distances rentrent dans cette catégorie telle que le mode de l'erreur (Ballester *et al.* 2003). Il est aussi possible de calculer l'histogramme de la valeur absolue de l'erreur (Wald *et al.* 1997) divisé par les valeurs des pixels de l'image de référence, ce qui permet d'obtenir la fonction de densité de probabilité.

Comme ces distances reposent sur l'image de différence, les valeurs idéales de chacune de ces distances sont 0, excepté pour la probabilité d'avoir une erreur plus petite qu'un seuil donné où la valeur idéale est alors 1.

2.4.2.3 Catégorie 3 : « corrélation »

Cette troisième catégorie regroupe les distances qui mesurent les similarités entre structures de petites tailles des images $B_k^*(i, j)$ et $B_k(i, j)$. Le coefficient de corrélation ($cc(B_k, B_k^*)$) est le plus célèbre d'entre eux. En effet, on a noté la profusion des publications y faisant référence : Blanc *et al.* (1998), Li (2000), Wald (2002), Ranchin *et al.* (2003), Rokhmatuloh *et al.* (2003), Gungor and Shan (2004), Alonso Reyes *et al.* (2005a), Gonzalez-Audicana *et al.* (2005), Otazu *et al.* (2005), Acerbi-Junior *et al.* (2006), Nencini *et al.* (2006), Thomas et Wald (2006b), Yang *et al.* (2006). Et cette liste n'est pas exhaustive. La valeur idéale du coefficient de corrélation est 1. Une distance équivalente est la covariance.

Wang & Bovik (2002) proposent un indice de qualité image qu'ils ont appelé *Q*. Rencontré aussi sous le nom de *UQI* pour *Universal Quality Index*, cet indice est une tentative pour prendre en compte la perception humaine et devient une référence dans le domaine de l'estimation de la qualité des produits fusionnés (Piella et Heijmans 2003, Alparone *et al.* 2004, Chen et Varshney 2005, Otazu *et al.* 2005, Zheng *et al.* 2005, Nencini *et al.* 2006, Thomas et Wald 2006b). Il résulte du produit de trois termes : le coefficient de corrélation, sensible aux hautes fréquences de l'image, un second terme relatif au contraste qui est fonction des écart-types des deux images, et enfin un terme de luminance

dépendant de la valeur du biais et donc des moyennes des deux images. L'expression de Q est donnée par :

$$Q_k(B_k, B_k^*) = f(L_k(B_k, B_k^*), C_k(B_k, B_k^*), S_k(B_k, B_k^*)) \quad \text{Eq. 2.7}$$

avec

$$Q_k(B_k, B_k^*) = cc(B_k, B_k^*) \frac{2M(B_k)M(B_k^*)}{M(B_k)^2 + M(B_k^*)^2} \frac{2\sigma(B_k)\sigma(B_k^*)}{\sigma(B_k)^2 + \sigma(B_k^*)^2} \quad \text{Eq. 2.8}$$

L'indice Q délivre des valeurs comprises dans l'intervalle $[-1, 1]$. La valeur idéale 1 est atteinte lorsque B_k^* coïncide parfaitement avec B_k .

Une amélioration, appelée $MSSIM_k$ (de l'acronyme anglais *Mean of Structure SIMilarity*) a été publiée par Wang *et al.* (2004). L'amélioration consiste à éviter les instabilités en prenant la précaution d'ajouter une constante aux dénominateurs de chaque terme afin de s'affranchir des divisions par zéro.

$$MSSIM_k(B_k, B_k^*) = \frac{\sigma(B_k, B_k^*) + \alpha_1}{\sigma(B_k)\sigma(B_k^*) + \alpha_1} \frac{2M(B_k)M(B_k^*) + \alpha_2}{M(B_k)^2 + M(B_k^*)^2 + \alpha_2} \frac{2\sigma(B_k)\sigma(B_k^*) + \alpha_3}{\sigma(B_k)^2 + \sigma(B_k^*)^2 + \alpha_3} \quad \text{Eq. 2.9}$$

Notons qu'une autre amélioration de la distance Q a été proposée par Piella & Heijmans (2003).

Lillo-Saavedra *et al.* (2005) ont utilisé un indice qu'ils ont appelé le *ERGAS* spatial. *ERGAS* est une distance globale qui sera expliquée dans la catégorie 7. Cette distance « *ERGAS* spatial » consiste à calculer tout d'abord la *RMSE* entre chaque modalité fusionnée Pan et le Pan après ajustement de la dynamique de ce dernier sur celle de la modalité MS. Ensuite, chaque *RMSE* est divisée par la moyenne de la modalité et les résultats sont sommés de manière quadratique, selon la formule de l'indice *ERGAS*. Nous ne recommandons pas d'utiliser cette distance parce que Pan ne peut en aucun cas représenter une référence valide pour les modalités MS fusionnées.

2.4.2.4 Catégorie 4 : « distances entre petites structures »

Cette catégorie ressemble à la catégorie précédente, puisqu'il s'agit aussi de mesurer la similarités entre les hautes fréquences des images. Mais cette fois-ci, une étape préalable est nécessaire puisqu'une extraction de ces hautes fréquences est effectuée au moyen de filtres passe-haut.

Zhou *et al.* (1998) et Li (2000) procède à l'extraction des hautes fréquences spatiales par convolution d'un filtre Laplacien convolué à la modalité fusionnée et à la modalité Pan avant de calculer la valeur du coefficient de corrélation entre les images obtenues. Cette quantité a aussi été retrouvée dans la littérature sous le nom de coefficient de corrélation spatial *sCC*. Cependant, comme nous l'avons souligné à de multiples reprises, ces auteurs n'appliquent pas correctement cette distance puisque qu'ils prennent la modalité Pan comme référence. Ce défaut a été corrigé par Otazu *et al.* (2005), nous appellerons $ccHF(B_k, B_k^*)$ cette distance. La valeur idéale du coefficient de corrélation est 1.

Eskicioglu et Fisher (1995), Li *et al.* (2001), Zheng *et al.* (2005) et Nencini *et al.* (2006) ont calculé le gradient pour chaque pixel B_k^* et B_k et ont moyenné la valeur obtenue sur tous les pixels de chaque image B_k^* and B_k . La distance est la différence entre les gradients moyens des deux images. La différence est donnée en relatif par rapport au gradient moyen de l'image de référence B_k . La valeur idéale de ce paramètre est 0.

Alonso-Reyes *et al.* (2005b) extraient le premier plan de coefficients d'ondelettes des images B_k^* et B_k et construisent un corrélogramme. Dans la situation idéale, les points du corrélogramme devraient se situer le long de la première diagonale ($y = x$). Une régression linéaire est appliquée sur le nuage de points et les paramètres qui donnent l'équation de la droite obtenue, sont comparés à ceux de la droite idéale ; c'est à dire une pente de valeur 1 et 0 comme ordonnée à l'origine.

Shi *et al.* (2004) proposent de comparer visuellement une ligne image avant et après fusion. Ils préconisent de particulièrement s'attarder sur les hautes fréquences qui ont été introduites lors de la fusion. Bien sûr, cette manière de procéder est tout à fait empirique et ne permet pas de quantifier

l'apport et la véracité des hautes fréquences introduites, de même qu'il n'est pas possible d'en retirer une idée générale sur la distorsion, c'est-à-dire la modification des basses fréquences de ce profil. Cette approche ne peut donc être retenue.

Vrabel (1996, 2000) définit une quantité qui témoigne de la résolution atteinte suite à un algorithme de fusion : c'est la *Ground Sampled Distance* (*GSD*, soit, en français, pas d'échantillonnage au sol). Il s'agit d'une loi mathématique définie de manière empirique. Nous n'avons pas sélectionné cette quantité car l'étude menée par le gouvernement américain dans IRARS (1995) pour la définition de l'échelle MIIRS a illustré le fait que deux images possédant des caractéristiques inverses (une possédant un contraste élevé avec un bruit faible et inversement) pouvaient présenter une *GSD* identique. La *GSD* n'est donc pas capable de donner une idée du bruit présent dans l'image. De plus, la *GSD* fait appel à un groupe d'interprètes qui donnent chacun une valeur de la *GSD* et la valeur finale de la *GSD* correspond à la moyenne des valeurs rapportées par le groupe d'interprètes. La *GSD* est donc très subjective et ne peut donner une prédiction fiable de la résolution atteinte. Notons de plus qu'il ne peut s'agir que d'une *GSD* virtuelle, correspondant à un pas d'échantillonnage équivalent ou virtuel, car la véritable définition du pas d'échantillonnage au sol est la projection du pixel de l'image sur le sol.

Nous attirons aussi l'attention sur le fait que les résultats obtenus par les distances de cette catégorie dépendent fortement de l'opérateur choisi pour extraire les hautes fréquences. Ainsi, si une méthode de fusion utilise un filtre Laplacien et que le même filtre Laplacien est utilisé dans la phase d'évaluation de la qualité, le résultat risque d'être biaisé.

2.4.2.5 Catégorie 5 : « mesure du gradient local »

La catégorie 5 traite des différences qui concernent la netteté des contours et des structures d'une image. Ces outils sont basés sur des grandeurs qui caractérisent la qualité des capteurs. Par exemple, une distance reposant sur la théorie de la *PSF* a été proposée (Li 2000, Lei et Tiziani 1989). La *PSF* est la réponse d'un capteur à une parfaite impulsion lumineuse que l'on appelle Dirac. La *PSF* révèle la manière dont le système, qu'il soit optique, acoustique ou autre, atténue le contraste de l'image en fonction de la fréquence spatiale. Li (2000) pose l'hypothèse que la *PSF* est une Gaussienne et définit le coefficient de flou (*Blur Parameter*) comme l'écart-type de cette Gaussienne. La distance proposée est la différence entre le paramètre de flou calculé pour l'image synthétisée et celui de l'image de référence.

Thomas et Wald (2006b) ont travaillé sur une amélioration de cet indice en proposant une distance qui repose sur la Fonction de Transfert de Modulation (*FTM*). La *FTM* est la transformée de Fourier de la *PSF*. La *FTM* peut être obtenue à partir des contours d'une image. Cette mesure est une courbe donnée dans le domaine fréquentiel qui informe sur les fréquences contenues dans le contour traité. Cet outil est utilisé pour mesurer la différence entre les fréquences contenues dans le contour de référence et pour ce même contour dans l'image synthétisée. Une norme L_2 est utilisée pour caractériser la distance entre ces deux courbes (L_2CFC). Les travaux sur cet outil basé sur l'estimation de la *FTM* sont expliqués dans le chapitre 4 de cette thèse.

Comme ces distances traitent des différences entre grandeurs appliquées à deux images, leur valeur idéale est 0.

2.4.2.6 Catégorie 6 : « distances entre signatures spectrales »

Dans cette catégorie 6, les distances visent à mesurer l'écart entre les vecteurs spectraux de l'image fusionnée et de référence. Ces distances quantifient l'erreur commise d'un point de vue spectral, ce qui est d'une importance déterminante pour des post-traitements de type classification.

Wald *et al.* (1997) proposent de compter le nombre de spectres qui concordent dans les deux images. Idéalement, ces deux nombres devraient être identiques, c'est-à-dire que leur différence devrait être nulle. Il est aussi possible de s'intéresser uniquement aux vecteurs spectraux les plus

fréquents puisque ce sont ceux qui auront le plus d'influence lors d'un processus de classification. Les écarts peuvent être quantifiés en comptant par exemple le nombre de spectres coïncidants et de comparer le nombre de pixels qui portent ces valeurs de spectre, ce qui correspond aux occurrences de ces vecteurs.

La distance *SAM* pour *Spectral Angle Mapper* commence à devenir un standard dans le domaine (Alparone *et al.* 2003, Park et Kang 2004, Nencini *et al.* 2006, Thomas et Wald 2006b). *SAM* est la valeur absolue de l'angle moyen entre les vecteurs spectraux de l'image fusionnée et de la référence. La valeur idéale du *SAM* est 0. *SAM* est donné par :

$$SAM(B(i, j), B^*(i, j)) = \arccos \left(\frac{\langle B^*(i, j), B(i, j) \rangle}{\|B^*(i, j)\|_2 \cdot \|B(i, j)\|_2} \right) \quad \text{Eq. 2.10}$$

A cette distance, Alparone *et al.* (2003) ont associé le critère de qualité suivant : il est désirable qu'un tel angle soit inférieur à celui mesuré entre le vecteur des données originales MS basse résolution ayant été rééchantillonnées à la haute résolution, et celui des données MS réelles à haute résolution spatiale.

$$SAM(B_0(i, j), B_0^*(i, j)) < SAM((B_1)^{interp}_0(i, j), B_0^*(i, j)) \quad \text{Eq. 2.11}$$

Thomas *et al.* (2005) calculent pour chaque pixel la différence de normes entre ces vecteurs spectraux. Ils en tirent deux distances qui sont le biais relatif ainsi que l'écart-type relatif à la moyenne des normes des vecteurs de la référence. Leur valeur idéale est 0.

$$diffNorms(i, j) = \|B(i, j)\| - \|B^*(i, j)\| \quad \text{Eq. 2.12}$$

Ces auteurs indiquent aussi qu'il est possible d'utiliser une autre distance basée sur la norme du vecteur résultant définie par la norme du vecteur correspondant à la différence entre un vecteur fusionné et un vecteur de référence au pixel (i, j) . Une image de norme des vecteurs résultants à chaque pixel est obtenue dont il est possible de donner la moyenne et l'écart-type. Les valeurs idéales sont 0, et l'expression de la norme du vecteur résultant est :

$$V_{res}(i, j) = \|B(i, j) - B^*(i, j)\| \quad \text{Eq. 2.13}$$

2.4.2.7 Catégorie 7 : « distances globales »

Cette catégorie regroupe les distances qui tentent, par une unique mesure, de donner une idée de la qualité globale pour l'ensemble B_0^* des modalités MS fusionnées. Munechika *et al.* (1993) ont, par exemple, proposé de sommer les *RMSE* obtenues pour chaque modalité MS tandis que Wald *et al.* (1997) en proposent une moyenne. Une première amélioration de cette distance a été définie ultérieurement par Wald (2000, 2002) : il s'agit de la distance *RASE* pour *Relative Average Spectral Error*. Son écriture est la suivante :

$$RASE(B_k, B_k^*) = 100/M \sqrt{\frac{1}{N} \sum_{k=1}^N RMSE(B_k, B_k^*)^2} \quad \text{Eq. 2.14}$$

où M correspond à la moyenne des N modalités MS de référence. *RASE* renvoie une valeur en pourcents dont la valeur idéale est 0. Elle caractérise la performance moyenne d'une méthode dans les modalités considérées. On remarque qu'aucun réajustement vis-à-vis de l'échelle n'est appliqué, ce qui rend cette mesure sensible à la résolution spatiale, ainsi qu'au ratio entre les images originales. Ces inconvénients ont été résolus avec le développement d'une nouvelle mesure nommée *ERGAS* qui est l'abréviation de *Erreur Relative Globale Adimensionnelle de Synthèse* (Wald 2000, 2002). Son calcul repose sur l'expression suivante :

$$ERGAS(B_k, B_k^*) = 100/ratio \sqrt{\frac{1}{N} \sum_{k=1}^N \left(\frac{RMSE(B_k, B_k^*)}{M(B_k)} \right)^2} \quad \text{Eq. 2.15}$$

Cette mesure est très souvent utilisée (Ranchin et Wald 2000, Ranchin *et al.* 2003, Ranchin *et al.* 2003, Gonzalez-Audicana *et al.* 2005, Otazu *et al.* 2005, Nencini *et al.* 2006, Thomas et Wald 2006b) car elle a la particularité d'être indépendante du ratio entre les images, du nombre de modalités

traitées, de la résolution spatiale, de la dynamique, du gain, des coefficients d'étalonnage et enfin des unités. Sa valeur idéale est 0.

Q_4 et Q_w correspondent à des extensions multimodales du coefficient Q présenté dans la catégorie 3. Q_4 est une distance qui s'appuie sur la théorie des quaternions pour mesurer la distance entre quatre modalités fusionnées avec leurs références (Alparone *et al.* 2004; Nencini *et al.* 2006). Q_w se calcule en moyennant les valeurs de Q obtenues pour chaque modalité MS (Otazu *et al.* 2005). La valeur idéale est 0 pour ces deux distances.

2.4.3 Redondances/complémentarités entre distances

Chaque catégorie de distances quantifie un aspect particulier de la qualité des produits de fusion. Dans cette section, nous montrons la complémentarité ou les redondances de certaines distances. La redondance permet de restreindre le nombre de distances à utiliser alors que la complémentarité est l'assurance d'un bilan de qualité complet. Les problèmes de l'évaluation de la qualité ne sont pas propres à la fusion d'images, mais concernent le domaine du traitement d'images en général dès lors qu'une image a été synthétisée et doit être comparée à sa référence. Nous allons montrer les risques à n'utiliser qu'une seule distance pour connaître la qualité d'une image.

2.4.3.1 Quelques illustrations simples

Le premier exemple placé en figure 2.2 concerne les distances de la catégorie 1. La figure 2.2a représente un bâtiment entouré de champs en périphérie de la ville de Starckville, aux Etats-Unis. Il s'agit de la modalité Pan du satellite Quickbird. A droite, la figure 2.2b correspond à un bruit gaussien que nous avons synthétisé de telle sorte que sa moyenne et son écart-type coïncident avec ceux de l'image de gauche.



Figure 2.2 : a) modalité Pan (Quickbird) - Starckville (USA), b) simulation sous IDL d'un bruit Gaussien. Copyright Digital Globe 2004.

Nous avons calculé quelques indices statistiques entre ces deux images. Ainsi, la différence entre les moyennes, entre les variances ou encore entre les entropies est nulle. Par contre, l'écart-type de l'image de différence relative à la moyenne de l'image de gauche et exprimée en pourcents atteint 42 %. Le corrélation et l'indice Q ont une valeur pratiquement égale à zéro. De plus, ces images n'ont pas le même nombre de spectres, ni de spectres les plus fréquents. Ces nombres montrent qu'il faut prendre garde à ne pas utiliser seulement un biais et/ou une différence de variances pour caractériser la différence entre deux images. Ici, on voit que les distances de la catégorie 3 « corrélation » sont sensibles à une telle différence, tout comme les distances de la catégorie 2 qui exploitent l'image de différence.

La figure 2.3 est un nouvel exemple : l'image de gauche a été acquise en mode panchromatique par satellite Ikonos qui montre un échangeur d'autoroutes situé à Fredericton au

Canada. L'image 2.3b correspond à la même image mais ayant subi une rotation de 180 degrés. Le contenu informatif de ces deux images est donc parfaitement identique.

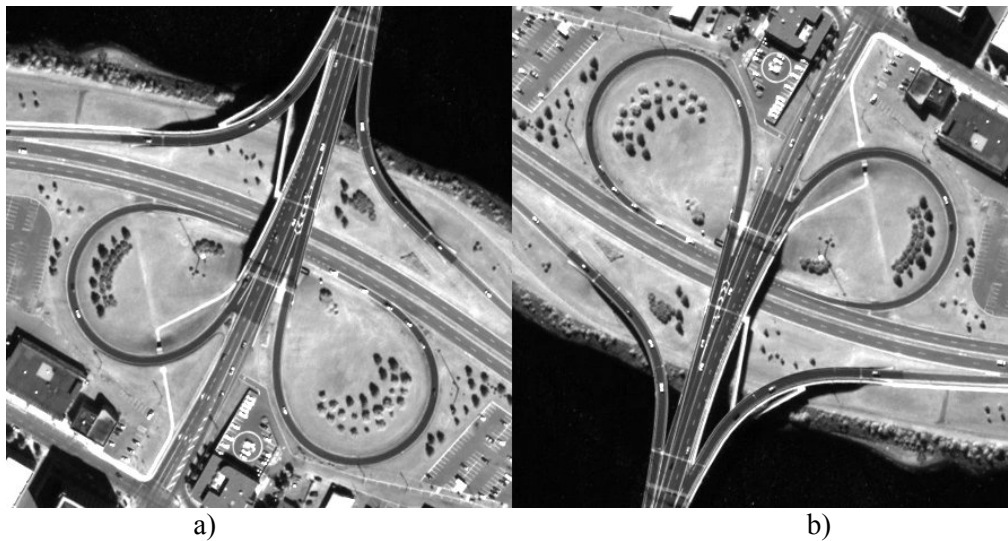


Figure 2.3 : a) modalité Pan (Ikonos) – Fredericton (Canada), b) même image ayant subi une rotation de 180 degrés. Copyright Space Imaging 2002.

Puisque ces deux images sont parfaitement identiques, les distances de la catégorie 1 sont parfaitement insensibles, alors que celles des catégories 2 et 3 sont très éloignées de leur valeur idéale. Par exemple, l'écart-type de l'image de différence est de 68 %, et l'entropie de l'image de différence a une valeur de 7. Ces distances sont donc sensibles aux différences de localisation des pixels dans une image, et seront donc sensibles à toute rotation ou translation. Finalement, lorsque le contenu informatif entre deux images est proche, les distances de la catégorie 1 donnent de bons résultats, c'est-à-dire des valeurs proches de leur valeur idéale.

La figure 2.4 va nous permettre de mettre en évidence le danger de n'utiliser comme distance uniquement le coefficient de corrélation. En temps que standard, le coefficient de corrélation est souvent le premier indice auquel on pense lorsque l'on entreprend une démarche d'évaluation de la qualité. Cette figure permet d'illustrer le fait que le coefficient de corrélation est insensible à une transformation de type affine $y = ax + b$. La figure 2.4a, acquise par le satellite Ikonos représente des parcelles agricoles situées dans la banlieue de Hasselt, une ville belge. L'image de droite est la même image dont tous les comptes numériques ont été multipliés par deux.

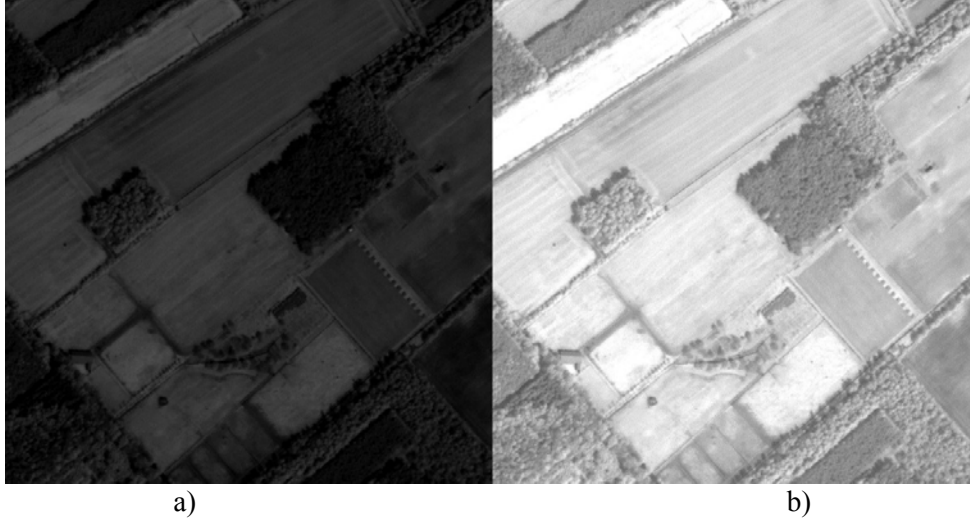


Figure 2.4 : a) Parcelles agricoles acquises par la modalité Pan (Ikonos) – Hasselt (Belgique), b) même modalité dont tous les comptes numériques ont été multipliés par deux. Copyright Space Imaging 2002.

Les résultats statistiques sont les suivants : le biais relatif à la moyenne de l'image de gauche est de -100% , la différence de variance relative atteint -300% et enfin l'écart type de l'image de différence donne une valeur de 12.6% . Par contre, le coefficient de corrélation atteint sa valeur idéale de 1, et le coefficient Q qui fait pourtant partie de la même catégorie, n'excède pas 0.6.

Ces quelques illustrations nous permettent d'affirmer que le coefficient de corrélation ne peut être utilisé seul, tout comme les distances de la catégorie 1. On remarque que le biais et la différence de variance réagissent lorsque le coefficient de corrélation reste insensible. Ces distances sont donc complémentaires. De plus, l'écart-type ou l'entropie informe sur l'énergie contenue dans l'image de différence. Au vu de ces résultats statistiques et de notre expérience, nous avons choisi d'utiliser les distances monomodales suivantes :

- ◆ le biais relatif ($biasRel(B_k, B_k^*)$),
- ◆ la différence en variances relative ($diffVarRel(B_k, B_k^*)$),
- ◆ l'écart-type de l'image de différence ($\sigma Rel(B_k, B_k^*)$),
- ◆ enfin, le coefficient de corrélation ($cc(B_k, B_k^*)$).

2.4.3.2 Relations mathématiques

La distance $RMSE$ est liée au biais et à l'écart-type par la relation suivante :

$$RMSE^2(B_k, B_k^*) = bias^2(B_k, B_k^*) + \sigma^2(B_k, B_k^*) \quad \text{Eq. 2.16}$$

Ainsi, si le biais est faible, il faut calculer soit $RMSE(B_k, B_k^*)$ soit $\sigma(B_k, B_k^*)$, sous peine d'obtenir un résultat redondant.

Les deux distances de la catégorie 5 (PSF et FTM) sont reliées par la forme suivante :

$$FTM = TF(PSF) \quad \text{Eq. 2.17}$$

où TF représente la transformée de Fourier. Inutile une fois encore de calculer ces deux distances pour caractériser la qualité des contours des images.

Enfin, on remarque que les trois indices SAM , $diffNorms$ et la norme de V_{res} sont reliées par le théorème d'Al Kashi, mieux connu sous le nom de théorème de Pythagore généralisé. Il s'écrit :

$$\|V_{res}(i, j)\|^2 = \|B(i, j)\|^2 + \|B^*(i, j)\|^2 - 2\|B(i, j)\| \|B^*(i, j)\| \cos(SAM) \quad \text{Eq. 2.18}$$

Ces distances ne sont donc pas complètement indépendantes, car si on forme :

$$diffNorms^2(i, j) = (\|B(i, j)\| - \|B^*(i, j)\|)^2 = \|B(i, j)\|^2 + \|B^*(i, j)\|^2 - 2\|B(i, j)\| \|B^*(i, j)\| \quad \text{Eq. 2.19}$$

La relation de Pythagore généralisée devient :

$$\|V_{res}(i, j)\|^2 = diffNorms^2(i, j) + 2\|B(i, j)\| \|B^*(i, j)\| (1 - \cos(SAM)) \quad \text{Eq. 2.20}$$

Le SAM devient progressivement une référence pour établir les bilans de qualité multimodaux (Park et Kang 2004) et nous recommandons d'utiliser cette distance. Pour compléter cette information, seule

l'une des deux autres grandeurs est nécessaire. Nous avons choisi d'étudier les statistiques liées à l'image de différence entre les normes des vecteurs spectraux.

2.4.3.3 Etude empirique

Nous avons exploré les relations qui existent entre le coefficient de corrélation et l'indice Q . En effet, l'indice Q est issue du produit de trois termes, dont l'un correspond au coefficient de corrélation. Nous avons procédé à une étude empirique sur quarante-six images qui provenaient de divers satellites (Quickbird, Ikonos, SPOT2, SPOT5). Les bandes spectrales des différentes modalités vont du bleu au PIR. Les ratios entre la haute et la basse résolution spatiale sont de 2 et 4. La méthode de fusion a arbitrairement été sélectionnée comme étant le modèle $M2$ proposé par Ranchin et Wald (2000). La particularité de cette méthode de fusion est de produire une image fusionnée dont la moyenne correspond à la moyenne des images de référence. Par conséquent, le biais est nul.

Pour chaque cas, les valeurs de l'indice Q ont été tracées en fonction des valeurs du coefficient de corrélation. La figure 2.5 montre les corrélogrammes obtenus pour les ratio 2 et 4. Une très forte similitude apparaît entre ces deux grandeurs mis à part quelques exceptions lorsque l'on se rapproche des corrélations plus faibles. La corrélation obtenue entre les deux séries de données est de respectivement 0.999 pour le ratio 2 et 0.989 pour le ratio 4. La valeur de la pente issue d'une régression linéaire est de 1.013 (resp. 0.878), et l'ordonnée à l'origine est de -0.012 (resp. 0.115).

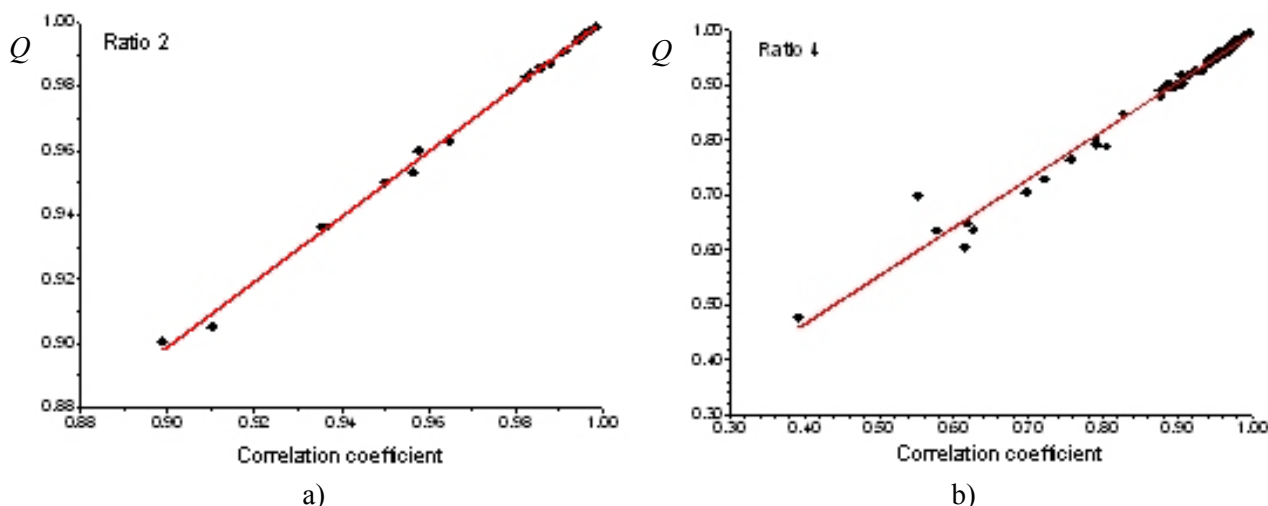


Figure 2.5 : a) corrélogramme entre Q et le coefficient de corrélation pour le ratio 2 (corrélation entre les deux séries : 0.999), b) idem pour le ratio 4 (corrélation entre les deux séries : 0.989)

La conclusion de notre étude est que, dans le cas d'une méthode de fusion qui conserve la moyenne comme les méthodes adoptant une approche multi-échelle et pour des valeurs de corrélation supérieure à 0.9, le coefficient de corrélation et Q donnent des résultats équivalents. Parce qu'il est plus simple à implanter, nous recommandons d'utiliser le coefficient de corrélation.

2.4.4 Proposition de bilans de qualité

L'interprétation des redondances et complémentarités entre distances a montré qu'obtenir un bilan de qualité complet d'un produit fusionné ou de manière générale d'une image synthétisée n'est pas une tâche aisée. Il faut tout d'abord bien connaître et formuler nos besoins ; en effet, un utilisateur cherchant à faire une classification sur ses images fusionnées privilégiera une bonne restitution des vecteurs multispectraux, alors que s'il cherche à faire une extraction des contours, il portera l'accent sur des distances davantage géométriques.

Les illustrations et relations mathématiques apportent un élément supplémentaire. Les distances qui appartiennent à des catégories différentes sont complémentaires et nous suggérons de sélectionner au moins une distance par catégorie. De plus, les distances regroupées dans une seule et même catégorie peuvent s'avérer soit complémentaires (exemple : le biais et la différence de variance) soit redondantes (exemple : la différence des variances, des écart-types ou des entropies). De plus,

comme le montre l'étude empirique sur le comportement des indices Q et cc , ils peuvent être complémentaires sur une certaine plage de valeurs et équivalent sur une autre (cc et $Q > 0.90$). Nous n'avons certainement pas relevé toutes les distances, ni exploré leurs interconnexions. Voilà pourquoi nous conseillons d'entreprendre toute démarche d'évaluation de la qualité avec précaution.

Cependant, au vu des résultats que nous avons obtenus et de nos expériences, nous avons pu sélectionner un certain nombre de distances. Si on cherche à obtenir un bilan de qualité monomodal, nous choisirions ces distances :

- ◆ le biais relatif de la catégorie 1 ($biasRel(B_k, B_k^*)$),
- ◆ la différence de variance relative appartenant aussi à la catégorie 1 ($diffVarRel(B_k, B_k^*)$),
- ◆ l'écart-type relatif de la catégorie 2 ($\sigma Rel(B_k, B_k^*)$),
- ◆ le coefficient de corrélation de la catégorie 3 ($cc(B_k, B_k^*)$),
- ◆ le coefficient de corrélation entre les hautes fréquences extraites par un filtre passe-haut de la catégorie 4 ($ccHF(B_k, B_k^*)$),
- ◆ et enfin la distance géométrique L_2CFC basée sur la FTM de la catégorie 5. Toutefois, ce dernier outil est très récent et nécessite des tests supplémentaires pour le valider.

Pour un bilan de qualité multimodal, nous utilisons deux des trois distances de la catégorie 6 : le SAM , le biais ($biasRel(diffNorms)$) et l'écart-type relatif ($\sigma Rel(diffNorms)$) de la différence des normes des vecteurs spectraux et la moyenne ($M(Vres)$) et l'écart-type ($\sigma(Vres)$) du vecteur résultant. Enfin, en ce qui concerne un bilan de qualité global, nous nous appuyerions principalement sur la distance $ERGAS$ (catégorie 7). Il est aussi tout à fait envisageable de former un bilan de qualité complet en analysant une ou plusieurs distances provenant d'une analyse monomodale et d'une analyse multimodale, comme par exemple, un coefficient de corrélation conjointement à l'indice SAM .

Les travaux futurs devraient maintenant porter sur des critères à appliquer à ces distances. Peu de publications ont concerné des critères pour définir si un produit de fusion est de bonne qualité ou non. Par exemple, l'étude de Wald (2002) sur $ERGAS$ a abouti à un seuil de 3 séparant bonne et mauvaise qualité ; elle a été faite avec des images de résolutions de type décimétrique. Des travaux similaires devraient être entrepris. Il nous semble que ces travaux sont par nature expérimentaux. Cependant les conclusions risquent de ne pas être faciles à généraliser. Cependant, ce seuil de 3 pour l'indice $ERGAS$ n'a absolument pas été vérifié lors de nos propres campagnes d'expérimentation, ni lors du concours IEEE sur la synthèse d'images (Alparone *et al.* 2006), qui portaient sur des images de résolutions de type décimétrique.

III Etude de l'hypothèse d'extrapolation du bilan de qualité

Des distances et bilans de qualité ont été recommandés pour la démarche de l'évaluation de la qualité des produits de fusion dans le chapitre précédent. Ce chapitre est consacré à la vérification de l'hypothèse d'extrapolation du bilan de qualité (BQ). Nous rappelons que cette hypothèse concerne la vérification de la propriété de synthèse. Comme aucune référence n'est disponible à la haute résolution spatiale pour comparer les produits de fusion, on suggère de se placer à une échelle de plus faible résolution. L'hypothèse d'extrapolation consiste à admettre que le BQ établi à *res1* est assez proche, voire moins bon que celui qui aurait pu être établi à la haute résolution *res0*, si le capteur MS considéré avait été conçu pour proposer la haute résolution spatiale. Or nous avons vu que, en théorie, cette hypothèse n'est pas vérifiée à coup sûr, mais qu'elle l'est parfois. En fait, on peut se poser les questions suivantes :

- ◆ Est-ce que cette hypothèse est souvent vérifiée ?
- ◆ Quel est son domaine de validité ?
- ◆ Est-il possible de prédire, à coup sûr, les cas pour lesquels l'hypothèse est ou non valide ?

Ce chapitre se structure de la manière suivante : une première partie 3.1 est consacrée à la description de la démarche suivie pour l'étude. La partie 3.2 liste les images et extraits sur lesquels nous avons procédé à l'étude empirique. Ensuite, la partie 3.3 indique les bilans de qualité choisis pour cette étude. Le paragraphe 3.4 présente des exemples de résultats obtenus pour deux extraits de paysage complètement différents en terme de signature spectrale et variabilité spatiale. Les résultats font l'objet de la partie 3.5 en proposant une approche progressive dans la vérification de l'hypothèse proposée. Leur interprétation est donnée dans la partie 3.6 qui offre des explications quant à l'origine de la vérification ou non de l'hypothèse.

3.1 Démarche suivie

Pour vérifier le protocole, l'idéal serait de posséder un ensemble MS fusionné B^* et de référence B , à basse résolution *res1* (bilan de qualité entre B_l et B_l^*) et à haute résolution spatiale *res0* (entre B_0 et B_0^*). Mais dans la réalité, cet ensemble B_0 n'existe pas. Pour observer l'évolution du BQ avec les échelles, nous avons été contraints de déplacer le problème dans les échelles, et de comparer les bilans aux résolutions *res1* et *res2*, et quand nous en avons la possibilité, entre des résolutions encore plus grossières (*res2* et *res3*, *res3* et *res4*).

Le déplacement du problème dans les échelles a nécessité le développement d'une méthode capable de synthétiser des images de référence à plusieurs résolutions pour les mêmes ensembles A et B placés en entrée, d'y appliquer ensuite une méthode de fusion et enfin de délivrer des BQ aux diverses échelles. L'outil mathématique qui nous permet d'atteindre des résolutions plus grossières en simulant des images de plus basse résolution spatiale est l'analyse multirésolution (Mallat 1989). A partir des résolutions initiales des images des ensembles A et B , on procède à une décomposition pyramidale hiérarchique de l'information en créant des versions de plus en plus grossières d'une image, appelées approximations successives, auxquelles on applique la méthode de fusion choisie. Ainsi, à chaque image fusionnée correspond une référence issue de cette décomposition hiérarchique. L'algorithme de décomposition hiérarchique de l'information que nous avons utilisé est la transformée en ondelettes « à trous » *ATWT*, acronyme pour *A Trous Wavelet Transform* (Dutilleux 1989). Cette transformée a pour caractéristique d'être dyadique, non-décimée et non-orthogonale.

L'objet de l'étude relative à ce chapitre n'est pas la méthode de fusion en elle-même, mais l'évolution du bilan de qualité dans les échelles pour une méthode de fusion donnée. Nous avons arbitrairement choisi la méthode *ATWT-M2* du concept *ARSIS* (Ranchin et Wald 2000). Cette méthode est dite 'globale' puisque la transformation des données extraites de l'ensemble *A* puis injectées dans *B* s'applique sur toute l'image, indépendamment de la région considérée. Elle est détaillée dans le chapitre 5 consacré à la description du concept et à ses principales implantations.

3.2 Images sélectionnées pour l'étude

Nous avons construit une base de données d'images offrant une bonne variabilité des paysages, aussi bien en terme de contenu spectral, que spatial (rivières, forêts, montagnes, cultures différentes, bâtiments, routes, ...). Cet échantillon de la population des images n'est certes pas suffisant pour servir de référence à toutes les tâches qui peuvent être accomplies en fusion d'images multispectrales, mais nous avons tenté de couvrir les paysages complexes ou moins complexes usuellement rencontrés par les utilisateurs d'images satellites.

Notre jeu de données comporte quatorze images satellites pour le ratio 2 et quinze pour le ratio 4. Les images satellites que nous avons sélectionnées ont été préalablement corrigées géométriquement. Les extraits sélectionnés sont de taille de 512 par 512 pixels. En ce qui concerne le ratio 2, nous avons travaillé avec des images provenant des satellites SPOT2 et SPOT5. Pour SPOT2, Il s'agit d'un extrait couvrant la ville de Barcelone avec une résolution spatiale de 20 m en mode multispectrale. Nous avons synthétisé des produits de fusion aux résolutions spatiales de 20, 40 et 80 m.

Du satellite SPOT5, nous disposons de deux sites géographiques : Marseille et Toulouse. La résolution spatiale des modalités multispectrales est de 10 m et nous avons été en mesure de produire des images fusionnées à 10, 20 et 40 m. Les extraits pour ces deux sites sont :

- ♦ Marseille : deux extraits qui couvrent le centre ville, deux des collines faiblement urbanisées, un des îles Frioul situées au large de Marseille, un des champs en périphérie de la ville d'Aubagne et enfin un dernier extrait en pleine mer Méditerranée,
- ♦ Toulouse : deux extraits du centre ville, deux extraits de parcelles agricoles dans la banlieue toulousaine, deux de quartiers résidentiels et une dernière qui permet de visualiser des entrepôts en zone industrielle.

Pour le ratio 4, nous avons travaillé sur de l'imagerie provenant des satellites Quickbird et Ikonos. Voici une synthèse des extraits sélectionnés. Pour le satellite Ikonos :

- ♦ Hasselt (Belgique) : deux extraits du centre ville, deux extraits comprenant des champs et des forêts et enfin un dernier extrait comprenant quelques habitations avec des jardins et quelques routes,
- ♦ Fredericton (Canada) : un extrait en centre ville, un extrait en pleine rivière

Les extraits Quickbird concernent les villes suivantes :

- ♦ Madrid (Espagne) : deux extraits du centre ville, n comprenant une forêt, un qui montre à la fois quelques arbres et des herbes rases et 1 extrait qui comprend quelques habitations avec leurs jardins et quelques chemins,
- ♦ Fredericton (Canada) : un extrait en centre ville, un extrait montrant des champs d'exploitation agricole et un extrait d'un échangeur autoroutier.

La modalité HR du satellite SPOT5 propose une résolution spatiale de 5 m. Ainsi, comme la résolution des modalités multispectrales est de 10 m, ce satellite offre l'opportunité de travailler avec un rapport de résolution de 2. Il existe une autre modalité de très haute résolution spatiale, appelée THR, est échantillonnée à 2.5 m. Mais la résolution spatiale annoncée de THR n'est qu'une approximation. En effet, cette modalité est particulière puisqu'elle est synthétisée à partir de

l'information provenant de deux barrettes CCD correspondant à une résolution de 5 m, placées l'une sous l'autre et décalées d'un demi-pixel. Par ailleurs, avant d'être mise à disposition, l'image THR a subi un échantillonnage en quinconce suivi d'un rééchantillonnage (Latry 2003). Les extraits du satellite SPOT5 auraient pu constituer autant de cas pour le ratio 4. Cependant, à cause de ces prétraitements, étant donné que l'image THR finale n'a pas été directement acquise par un capteur comme le sont les autres images de notre base de données image, nous avons préféré omettre dans la présentation l'ensemble de ces résultats. Notons toutefois que les tests avec ces images ne contredisent pas les résultats obtenus pour les autres images de ratio 4. La figure 3.1 suivante montre la diversité des images de notre étude.

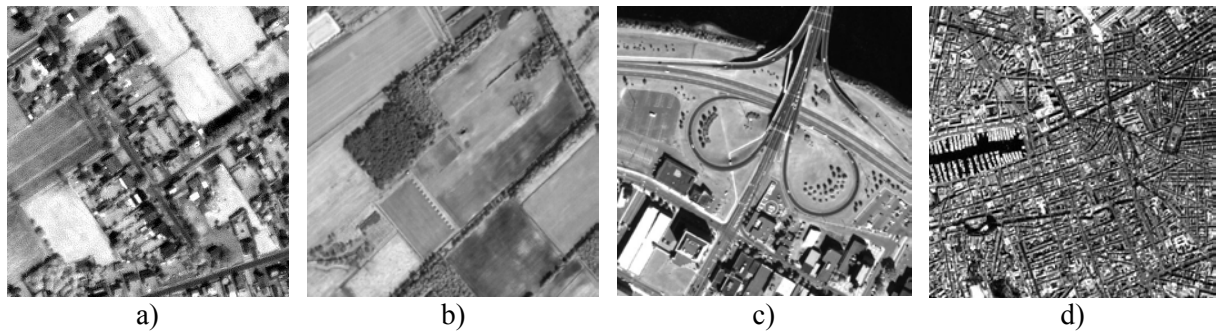


Figure 3.1 : diversité des images. a) satellite Ikonos – banlieue de Hasselt en Belgique, b) parcelles agricoles de la même zone géographique, c) échangeur autoroutier, Fredericton au Canada, Quickbird et d) centre ville de Marseille, SPOT5. Copyright respectifs pour Ikonos, Quickbird, et SPOT5 : Space Imaging 2002, Digital Globe 2002, CNES 2000.

3.3 Choix des bilans de qualité (BQ)

Nous distinguons trois catégories de bilans : les BQ monomodaux, multimodaux, et enfin globaux. Pour les BQ monomodaux, nous avons décidé d'étudier quatre BQ : le coefficient de corrélation (cc), l'indice Q , la réunion entre l'écart-type relatif et le coefficient de corrélation (σRel , cc) et enfin la réunion entre les deux précédents et la différence des variances relatives à la variance de la référence ($diffVarRel$, σRel , cc). Notons que dans le chapitre précédent, nous avons aussi recommandé d'utiliser comme distance monomodale le biais. Cependant, comme la méthode de fusion utilisée pour les tests repose sur l'injection de coefficient d'ondelettes dont le premier moment (la moyenne) est nulle, nous avons exclu cette distance qui n'est pas discriminante dans ce cas.

En ce qui concerne les BQ multimodaux, nous avons sélectionné l'indice SAM ainsi que les deux distances relatives au vecteur résultant, c'est-à-dire le vecteur résultant moyen sur tout l'image et son écart-type ($M(V_{res})$, $\sigma(V_{res})$). Ce choix est en accord avec les démonstrations mathématiques rapportées au chapitre précédent à savoir que l'indice SAM utilisé seul est incomplet et nécessite d'être complété soit par des informations sur les statistiques de la différence des normes, soit sur celles du vecteur résultant que nous avons privilégiées ici.

Et enfin, pour les bilans globaux, nous utilisons l' $ERGAS$ déjà exploité par la littérature, la réunion d'une distance monomodale et d'une distance multimodale (cc , SAM) et enfin le BQ que nous recommandons qui est la réunion entre : $diffVarRel$, σRel , cc , $ERGAS$, SAM , $M(V_{res})$, $\sigma(V_{res})$. Néanmoins, il manque le coefficient de corrélation entre les hautes fréquences extraites des images fusionnées et de référence car, au moment de cette étude, cette distance n'était pas implantée dans notre plate-forme qualité. Le tableau 3.1 synthétise les différents BQ choisis pour l'analyse.

<i>Indice du BQ</i>	<i>Type de bilan</i>	<i>Distances préconisées</i>
1	monomodal	$[cc]$
2	monomodal	$[Q]$
3	monomodal	$[\sigma Rel, cc]$
4	monomodal	$[diffVarRel, \sigma Rel, cc]$
5	multimodal	$[SAM]$
6	multimodal	$[M(Vres), \sigma(Vres)]$
7	global	$[ERGAS]$
8	global	$[cc, SAM]$
9	global	$[diffVarRel, \sigma Rel, cc, ERGAS, SAM, M(Vres), \sigma(Vres)]$

Tableau 3.1 : les bilans de qualité considérés pour la validation de l'hypothèse d'extrapolation.

La vérification de l'hypothèse consiste tout d'abord, à effectuer un bilan de qualité monomodal et multimodal par échelle (ou résolution) fusionnée, puis à étudier l'écart existant entre les bilans établis entre deux résolutions successives.

3.4 Exemples de résultats obtenus par la plate-forme qualité

La figure 3.2 correspond à l'extrait d'une image dans la région périphérique de la ville de Marseille, ville située au sud de la France. Le type de paysage que couvre cette image est une zone aride composée de collines recouvertes de garrigue, végétation typique des régions méditerranéennes. Cette image est la modalité Pan du satellite SPOT5, qui présente une résolution spatiale de 5 m au sol. La figure est associée aux tableaux 3.2 et 3.3 qui correspondent aux statistiques mono- et multimodales spécifiées dans la partie précédente. Ces statistiques sont délivrées pour les résolutions spatiales $res1 = 10$ m, $res2 = 20$ m, $res3 = 40$ m et enfin $res4 = 80$ m. Pour cette image, nous avons 4 modalités MS disponibles : la modalité MIR pour Moyen InfraRouge, la modalité PIR pour Proche InfraRouge, et les modalités rouge et verte.



Figure 3.2 : image SPOT5, composition colorée PIR, rouge et verte – collines de la région marseillaise
– Copyright CNES 2002.

Statistiques monomodales		<i>res1 = 10 m</i>	<i>res2 = 20 m</i>	<i>res3 = 40 m</i>	<i>res4 = 80 m</i>
<i>diffVarRel</i>	MIR	-0.5	-0.7	1.8	-0.4
	PIR	-0.9	-0.1	-0.2	-0.2
	rouge	-1.0	-0.1	0.0	-0.2
	vert	0.5	1.8	3.5	0.8
<i>σRel</i>	MIR	2.7	1.7	1.5	1.9
	PIR	1.5	0.5	0.5	0.7
	rouge	2.1	0.8	0.7	1.0
	vert	2.9	3.0	3.3	3.0
<i>cc</i>	MIR	0.995	0.997	0.998	0.997
	PIR	0.998	1.000	1.000	0.999
	rouge	0.997	1.000	1.000	0.999
	vert	0.991	0.988	0.981	0.990
<i>Q</i>	MIR	0.995	0.997	0.998	0.997
	PIR	0.998	1.000	1.000	0.999
	rouge	0.997	1.000	1.000	0.999
	vert	0.991	0.988	0.980	0.990

Tableau 3.2 : statistiques monomodales pour l'extrait de la figure 3.2.

<i>Statistiques multimodales</i>	<i>res1 = 10 m</i>	<i>res2 = 20 m</i>	<i>res3 = 40 m</i>	<i>res4 = 80 m</i>
<i>ERGAS</i>	1.17	0.93	0.90	0.94
<i>SAM</i>	0.71	0.78	0.80	0.87
<i>biaisRel(diffNorms)</i>	0.023	0.028	-0.026	-0.074
<i>σRel(diffNorms)</i>	2.0	1.3	1.2	1.1
<i>M(Vres)</i>	2.6	2.2	2.1	2.2
<i>σ(Vres)</i>	1.9	1.4	1.3	1.3

Tableau 3.3 : statistiques multimodales pour l'extrait de la figure 3.2.

La figure 3.3 est une image acquise au-dessus du centre ville de Madrid. L'image, acquise par la modalité Pan du satellite Quickbird (0.7 m de résolution spatiale), montre un grand nombre de routes et de rues plus étroites qui s'intersectent. Les tableaux 3.4 et 3.5 suivent la même logique que celle des tableaux précédents en montrant les résultats obtenus concernant les statistiques mono- et multimodales pour cette image. La fusion a été appliquée sur trois modalités MS dont les bandes spectrales sont situées dans le visible : rouge, verte et bleue. Les résolutions spatiales de synthèse sont 2.8 m et 11.2 m.



Figure 3.3 : centre ville de Madrid (Espagne), composition colorée rouge, verte et bleue – Quickbird – Copyright Digital Globe 2000.

<i>Statistiques monomodales</i>		<i>res1 = 2.8 m</i>	<i>res2 = 11.2 m</i>
<i>diffVarRel</i>	rouge	-16.8	3.9
	vert	-17.2	-3.1
	bleu	-17.9	5.6
<i>σRel</i>	rouge	5.4	1.3
	vert	5.3	3.2
	bleu	5.0	3.5
<i>cc</i>	rouge	0.984	0.996
	vert	0.970	0.962
	bleu	0.944	0.912
<i>Q</i>	rouge	0.981	0.996
	vert	0.967	0.962
	bleu	0.940	0.912

Tableau 3.4 : statistiques monomodales pour l'extrait de la figure 3.3.

<i>Statistiques multimodales</i>	<i>res1 = 2.8 m</i>	<i>res2 = 11.2 m</i>
<i>ERGAS</i>	1.31	0.73
<i>SAM_{moy}</i>	1.03	0.93
<i>biaisRel(diffNorms)</i>	-0.015	-0.387
<i>σRel(diffNorms)</i>	4.8	2.4
<i>M(Vres)</i>	19.1	10.9
<i>σ(Vres)</i>	14.9	8.4

Tableau 3.5 : statistiques multimodales pour l'extrait de la figure 3.3.

3.5 Résultats statistiques pour la vérification de l'hypothèse

3.5.1 Une vérification par cas

Nous avons d'abord étudié le cas où les BQ sont identiques ou assez proches aux différentes résolutions, puis le cas où le BQ à haute résolution est meilleur que celui à basse résolution. Ainsi, nous avons quatre cas :

- ◆ Cas 1 : les BQ sont identiques.
- ◆ Cas 2 : les BQ sont assez proches.
- ◆ Cas 3 : le BQ est meilleur à l'échelle la plus fine.
- ◆ Cas 4 : soit les BQ sont assez proches ou identiques, soit le BQ est meilleur à l'échelle la plus fine.

Pour chacun de ces cas, nous comptons le nombre d'images favorables.

La proximité entre deux distances a été définie par des critères. Tout d'abord, nous avons admis que deux distances sont identiques à partir du moment où les points reportés sur un graphique sont « visuellement superposés ». Nous avons calculé la distance maximale à partir de laquelle nous considérons que les deux points ne sont plus superposés. Ces valeurs sont rapportées dans la deuxième colonne du tableau 3.6. De la même façon, nous avons défini des seuils pour l'hypothèse « assez proches ». Les résultats sont placés dans la troisième colonne de ce même tableau. Les valeurs idéales sont rappelées pour chacune de ces distances dans la dernière colonne.

Distances	Intervalle de l'hypothèse « identiques »	Intervalle de l'hypothèse « assez proches »	Rappel des valeurs idéales
<i>diffVarRel</i>	+/- 0.3 %	+/- 2.5 %	0
<i>σRel</i>	+/- 0.02 %	+/- 2.5 %	0
<i>cc</i>	+/- 0.002	+/- 0.025	1
<i>Q</i>	+/- 0.002	+/- 0.025	1
<i>ERGAS</i>	+/- 0.05	+/- 0.5	0
<i>SAM</i>	+/- 0.03°	+/- 0.5°	0
<i>biaisRel(diffNorms)</i>	+/- 0.005 %	+/- 0.05 %	0
<i>σRel(diffNorms)</i>	+/- 0.1 %	+/- 2.5 %	0
<i>M(Vres)</i>	+/- 0.2	+/- 2.5	0
<i>σ(Vres)</i>	+/- 0.1	+/- 2.5	0

Tableau 3.6 : intervalles de validité des distances pour chaque cas de vérification de l'hypothèse.

Exemple : admettons que l'on considère le BQ correspondant au coefficient de corrélation. Nous avons dégradé une image de manière à créer des produits de fusion à *res1* et *res2*. Un coefficient de corrélation *cc* de 0.955 est obtenu à *res2*. Alors, si le coefficient de corrélation appartient à l'intervalle :

- ◆ $cc \in [0.953, 0.957] \Rightarrow$ les BQ sont considérés comme identiques,
- ◆ $cc \in [0.93, 0.98] \Rightarrow$ les BQ sont considérés comme assez proches,
- ◆ $cc \in [0.955, 1] \Rightarrow$ le BQ à *res1* est meilleur que celui à *res2*,
- ◆ $cc \in [0.93, 1] \Rightarrow$ l'hypothèse d'extrapolation est vérifiée.

On remarque que l'hypothèse d'extrapolation (cas 4) correspond à la réunion des cas 2 et 3. Ainsi, si le BQ est constitué d'une seule distance, comme le sont les bilans BQ1, BQ2, BQ5 et BQ7, alors le nombre d'images qui vérifie le cas 4 peut se déduire des deux précédents. Soit α l'ensemble des images vérifiant l'hypothèse du cas 2 et β celui du cas 3. Si on définit le nombre d'images pour ces deux cas par leur cardinal $\text{card}(\alpha)$ et $\text{card}(\beta)$, alors le cardinal pour le cas 4 correspondant à l'hypothèse d'extrapolation représente la somme des cardinaux à laquelle on a soustrait le cardinal de l'intersection :

$$\text{card}(\alpha \cup \beta) = \text{card}(\alpha) + \text{card}(\beta) - \text{card}(\alpha \cap \beta) \quad \text{Eq. 3.1}$$

3.5.2 Présentation de la forme des résultats

Les résultats sont récapitulés sous forme de tableaux, avec un tableau différent par BQ. Au départ, neuf BQ ont été sélectionnés. Néanmoins, seulement huit tableaux de résultats vont être présentés car le coefficient de corrélation et le coefficient *Q* ont donné des résultats équivalents avec une différence n'excédant jamais 5 % des images. La redondance des pourcentages ne présentait aucun intérêt pour notre interprétation. La structure générale d'un tableau de résultats est la suivante (tableau 3.7) :

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images				
Cas 1				
Cas 2				
Cas 3				
Cas 4				

Tableau 3.7 : forme du tableau de résultats pour un BQ donné

Les tableaux de résultats comportent deux colonnes principales qui distinguent les images de ratio 2 de celles de ratio 4. Pour le ratio 2, trois sous-colonnes permettent l'étude de l'évolution du BQ entre *res2* et *res1*, entre *res3* et *res2* et entre *res4* et *res3* lorsque nous en avons la possibilité. Pour le ratio 4, une seule colonne nous donne l'évolution du BQ entre *res2* et *res1*. Pour chacune de ces sous-colonnes, nous avons répertorié le nombre total de cas disponibles.

Chaque tableau comporte quatre lignes de résultats qui reprennent les quatre cas de la vérification de l'hypothèse d'extrapolation. Ces cas permettent d'aboutir à la dernière ligne, qui délivre véritablement les résultats concernant notre hypothèse de travail. Le nombre d'images recensé pour chacune des cas est donné en nombre d'images, suivi entre parenthèses du pourcentage par rapport au nombre total d'images.

3.5.3 Résultats pour les BQ monomodaux

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images	53	53	29	59
Cas 1	19 (36 %)	35 (66 %)	14 (48 %)	8 (14 %)
Cas 2	44 (83 %)	43 (81 %)	25 (86 %)	27 (46 %)
Cas 3	30 (57 %)	49 (92 %)	25 (86 %)	51 (86 %)
Cas 4	50 (94 %)	49 (92 %)	27 (93 %)	59 (100 %)

Tableau 3.8 : BQ1 = [cc], résultats à peu près équivalents pour BQ2.

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images	53	53	29	59
Cas 1	3 (6 %)	15 (28 %)	10 (34 %)	2 (3 %)
Cas 2	41 (77 %)	42 (79 %)	24 (83 %)	22 (37 %)
Cas 3	10 (19 %)	19 (36 %)	20 (69 %)	34 (58 %)
Cas 4	46 (87 %)	48 (91 %)	27 (93 %)	54 (92 %)

Tableau 3.9 : BQ3 = [σ_{Rel} , cc]

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images	53	53	29	59
Cas 1	3 (6 %)	15 (28 %)	3 (10 %)	0 (0 %)
Cas 2	39 (74 %)	42 (79 %)	15 (52 %)	4 (7 %)
Cas 3	10 (19 %)	18 (34 %)	20 (69 %)	34 (58 %)
Cas 4	46 (87 %)	48 (91 %)	26 (90 %)	47 (80 %)

Tableau 3.10 : BQ4 = [*diffVarRel*, σ_{Rel} , *cc*]

3.5.4 Résultats pour les BQ multimodaux

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images	16	16	9	17
Cas 1	4 (25 %)	4 (25 %)	3 (33 %)	4 (24 %)
Cas 2	14 (88 %)	16 (100 %)	8 (89 %)	11 (65 %)
Cas 3	13 (81 %)	8 (50 %)	5 (56 %)	12 (71 %)
Cas 4	14 (88 %)	16 (100 %)	9 (100 %)	16 (94 %)

Tableau 3.11 : BQ5 = [*SAM*]

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images	16	16	9	17
Cas 1	2 (13 %)	3 (19 %)	4 (44 %)	2 (12 %)
Cas 2	14 (88 %)	16 (100 %)	8 (89 %)	3 (18 %)
Cas 3	6 (38 %)	8 (50 %)	7 (78 %)	7 (41 %)
Cas 4	14 (100 %)	16 (100 %)	9 (100 %)	8 (47 %)

Tableau 3.12 : BQ6 = [*M(Vres)*, $\sigma(Vres)$]

3.5.5 Résultats pour les BQ globaux

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images	16	16	9	17
Cas 1	3 (19 %)	3 (19 %)	2 (22 %)	4 (24 %)
Cas 2	11 (69 %)	13 (81 %)	8 (89 %)	10 (59 %)
Cas 3	7 (44 %)	7 (44 %)	6 (67 %)	11 (65 %)
Cas 4	11 (69 %)	13 (81 %)	9 (100 %)	15 (88 %)

Tableau 3.13 : BQ7 = [*ERGAS*]

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images	53	53	29	59
Cas 1	5 (9 %)	11 (21 %)	7 (24 %)	4 (7 %)
Cas 2	38 (72 %)	44 (83 %)	24 (83 %)	22 (37 %)
Cas 3	28 (53 %)	24 (45 %)	19 (66 %)	41 (69 %)
Cas 4	48 (91 %)	49 (92 %)	27 (93 %)	57 (97 %)

Tableau 3.14 : BQ8 = [cc, SAM]

	Ratio 2			Ratio 4
	<i>res2-res1</i>	<i>res3-res2</i>	<i>res4-res3</i>	<i>res2-res1</i>
Nb total d'images	53	53	29	59
Cas 1	0 (0 %)	6 (11 %)	4 (14 %)	0 (0 %)
Cas 2	17 (32 %)	30 (57 %)	14 (48 %)	4 (7 %)
Cas 3	8 (15 %)	15 (28 %)	15 (52 %)	22 (37 %)
Cas 4	35 (66 %)	44 (83 %)	26 (90 %)	28 (47 %)

Tableau 3.15 : BQ9 = [diffVarRel, σ Rel, cc, ERGAS, SAM, $M(Vres)$, $\sigma(Vres)$]

3.6 Description et interprétation des résultats

3.6.1 Grandes tendances

La profusion des résultats a demandé quelques efforts dans leur interprétation. Une des grandes tendances tirée des différents tableaux est que les pourcentages ont tendance à augmenter en avançant dans les cas. En effet, les pourcentages les plus élevés sont situés en dernière ligne des tableaux correspondant au cas 4, c'est-à-dire la vérification de l'hypothèse d'extrapolation. Excepté pour le dernier BQ qui est le plus contraignant, les pourcentages des dernières lignes de tous les tableaux sont supérieurs à 70 %, ce qui signifie qu'une première conclusion de cette étude est que l'hypothèse d'extrapolation est souvent vérifiée.

Une autre conclusion générale est que les valeurs des pourcentages sont très dépendantes du BQ considéré. Plus le BQ contient de distances, plus celui-ci aura tendance à présenter des pourcentages plus faibles.

Pour le ratio 2, les pourcentages ont tendance à augmenter avec les échelles en passant de l'étude à *res2-res1* à *res3-res2*, mais il est très difficile de conclure pour les échelles plus grossières (*res3-res2* à *res4-res3*) où, là encore, les conclusions dépendent fortement du BQ. Cependant, il faut prendre garde à la signification des pourcentages rapportés vis-à-vis du nombre total d'images dans nos conclusions. En effet, lorsque vingt-neuf images sont disponibles, on incrémente de 3 % à chaque fois qu'une image supplémentaire vérifie un cas, alors qu'on n'augmente que d'un ou deux pourcents lorsque le nombre total d'images s'élèvent à cinquante trois.

Une analyse plus fine des résultats nous permet de comprendre le comportement des différents BQ vis-à-vis des images testées et des différents cas de vérification de l'hypothèse.

3.6.2 Description des tableaux de résultats

Chaque cas est rappelé en gras et en italique. Pour faciliter la compréhension, on se base sur la description du BQ correspondant à l'indice *ERGAS* car la littérature a déjà montré l'intérêt de cet indice (Wald 2002). Ensuite, nos conclusions sont élargies avec l'analyse des autres tableaux.

Cas 1 : Les BQ sont identiques

La première ligne, première colonne, de résultats du tableau *ERGAS* montre que trois images sur les seize disponibles délivrent un *ERGAS* à *res2* identique à celui obtenu à *res1*, soit 19 % des images. On observe que le pourcentage d'images est identique à celui entre *res3* et *res2*, et est équivalent à celui de 22 % obtenu entre *res4* et *res3* parce que le nombre total d'images a été divisé de moitié (neuf images). Pour le ratio 4, 24 % des images répondent favorablement à ce premier cas. Cela signifie que, quel que soit le ratio ou la résolution, environ 80 % des images montrent des *ERGAS* qui ne sont pas identiques aux diverses échelles que nous avons synthétisées.

Comme nous l'avons annoncé dans les grandes tendances, les résultats fluctuent fortement en fonction du BQ. Ainsi, même si les résultats pour le cas 1 sont les plus faibles pour chaque tableau, ils peuvent avoir des valeurs moyennes de l'ordre de quelques pourcents pour le BQ9 à 30 % pour la distance *SAM* qui forme le BQ5. Ceci est en accord avec la remarque sur le nombre de distances qui composent les BQ. En effet, le BQ9 contient sept distances, ce qui minimise la probabilité que toutes les distances valident simultanément la même hypothèse. Les BQ1, 5 et 7 ne sont formés que d'une seule distance et sont donc les moins contraignants. Leurs pourcentages pour ce premier cas sont situés entre 20 % et 30 % des images de la base de données.

En ce qui concerne le ratio 2, les pourcentages de la première colonne qui correspondent aux résolutions spatiales les plus fines sont les plus faibles. Cette conclusion est d'ailleurs valable pour pratiquement toutes les lignes de tous les tableaux. Par contre, le pourcentage de la colonne 2 est plus faible ou équivalent à celui de la colonne 3 pour les BQ3, 5, 6, 7 et 8 mais la conclusion est inversée pour les autres BQ.

Pour le cas 1, si nous comparons les résultats obtenus pour les ratios 2 et 4 entre *res2* et *res1*, les pourcentages sont équivalents pour chaque BQ.

Cas 2 : Les BQ sont assez proches

On relâche la première hypothèse en appliquant un intervalle d'acceptabilité plus large autour de chaque distance grâce aux valeurs définies dans le tableau 3.6. La seconde ligne du tableau du BQ contenant la distance *ERGAS* montre un saut important des pourcentages. Des résolutions spatiales les plus fines vers les plus grossières, les valeurs des pourcentages sont de 69 %, 81 % et 89 %. Par conséquent, environ trois images sur quatre valident l'hypothèse que les *ERGAS* sont « assez proches » aux différentes échelles étudiées. Le ratio 4 donne une valeur légèrement moins élevée de 59 % des images, mais fait tout de même pencher la balance en faveur de l'hypothèse puisque le pourcentage excède 50 %.

L'amplification des valeurs des pourcentages se généralise aux autres tableaux. Pour le ratio 2, les valeurs des pourcentages des BQ1, 2, 3, 5, 6 et 8 avoisinent ou dépassent 80 %. La troisième colonne du BQ4 est plus faible que les deux autres mais son pourcentage dépasse 50 %. Les valeurs du BQ9 sont plus faibles : un tiers des images sont favorables lorsque la résolution spatiale est élevée, le pourcentage se rapproche des 60 % pour les échelles intermédiaires pour finalement chuter à 48 % à la basse résolution.

Par contre, pour le ratio 4, les valeurs sont moins élevées. Les BQ contenant le plus de distances (BQ4 et BQ9) donnent seulement un pourcentage de 7 %. Les BQ contenant deux distances (BQ3, 6 et 8) ont des pourcentages entre 20 et 40 % pour atteindre et même dépasser les 50 % pour les BQ avec une seule distance.

Cas 3 : le BQ est meilleur à l'échelle plus fine

Il s'agit maintenant de vérifier le nombre d'images qui valident le second aspect de l'hypothèse d'extrapolation. Si on s'appuie sur le tableau de *ERGAS*, le ratio 2 montre des valeurs plus faibles que pour le cas précédent, ce qui n'est pas vérifié pour le ratio 4. Les valeurs sont identiques entre *res2-res1* et *res3-res2* avec une valeur de 44 %, avec un accroissement de la valeur lors de la comparaison entre BQ à *res3-res4* avec un pourcentage de 67 %. En ce qui concerne le ratio 4, les valeurs des pourcentages augmentent en s'approchant de l'hypothèse d'extrapolation, avec pour ce troisième cas, une valeur de 65 %.

Les autres BQ montrent eux aussi un accroissement régulier des pourcentages en ce qui concerne le ratio 4. Tous ces pourcentages dépassent 50 % excepté pour les BQ6 et BQ9, où les pourcentages ne sont qu'à 40 % des images. Pour le ratio 2, les valeurs ont tendance à augmenter en décroissant la résolution. Cette observation est valable pour les BQ3, 4, 5, 6, 7 et 9. Par contre, comme le nombre d'images influence fortement les pourcentages, on peut admettre que les pourcentages des BQ1 et 2 pour *res3-res2* et *res4-res3* sont équivalents. Le BQ8 offre un pourcentage un peu plus élevé à haute résolution, mais on peut conclure que les résultats sont comparables, quel que soit le BQ.

Cas 4 : soit les BQ sont assez proches, soit le BQ est meilleur à l'échelle plus fine

La réunion des 'sous-hypothèses' précédentes (cas 2 et 3) aboutit à l'hypothèse d'extrapolation. Les résultats correspondent aux valeurs des statistiques situées dans la dernière ligne des tableaux pour chaque BQ. Comme la théorie nous l'indique lorsqu'une seule distance constitue le BQ, le cardinal de la réunion de deux ensembles se calcule en ajoutant les cardinaux des deux ensembles qui la composent, sans oublier de soustraire le cardinal de leur intersection. Les pourcentages obtenus ont des valeurs comprises entre la somme des deux pourcentages trouvés pour les deux sous-hypothèses précédentes, soit être égaux à l'un d'entre eux. L'identité avec l'un des pourcentage n'est possible que si l'un des sous-ensembles contient le second. Les résultats des tests témoignent que nous sommes en accord avec la théorie puisque les BQ1, BQ2, BQ5 et BQ7 vérifient ces remarques.

La dernière ligne du tableau *ERGAS* présente des valeurs supérieures ou égales à celles contenues dans leur colonne. On remarque qu'on obtient les mêmes pourcentages aux cas 2 et 4 pour les résolutions *res2-res1* et *res3-res2* du ratio 2. Ces résultats signifient que, dans ces deux situations, le cardinal de l'hypothèse du cas 3 est inclus dans celui du cas 2. Les pourcentages augmentent en diminuant la résolution spatiale pour atteindre 100 % des images pour une comparaison des *ERGAS* entre *res4-res3*. Le pourcentage vaut presque 90 % des images pour le ratio 4, ce qui signifie que pour la plupart des images de la base de données images, l'hypothèse est vérifiée.

Les autres tableaux de résultats confirment l'idée que les pourcentages des dernières lignes majorent ceux de chaque tableau par colonne. En ce qui concerne le ratio 2, les pourcentages sont très élevés et donc plutôt en faveur de l'hypothèse d'extrapolation. Tous les pourcentages atteignent 100 %, excepté pour la haute résolution des BQ7 et 9 où la valeur du pourcentage n'est que de 70 % environ. Plusieurs tableaux présentent des résultats très proches de 100 % des images.

Pour le ratio 4, les résultats sont aussi favorables à l'hypothèse d'extrapolation avec sept BQ sur les neuf présentant un pourcentage supérieur à 80 %. Les résultats sont beaucoup plus mitigés pour le BQ6 qui correspond aux statistiques du vecteur résultant et pour le BQ9 qui correspond à la réunion de toutes les distances où le résultat est à 50 % des images. Pour ces deux BQ, aucune des deux hypothèses, que ce soit l'hypothèse d'extrapolation ou son contraire, n'est privilégiée. On remarque qu'à part pour ces deux cas, les conclusions sur l'hypothèse d'extrapolation ne sont pas sensibles au changement de ratio.

En conclusion, nous pouvons donc affirmer qu'avec les critères que nous nous sommes fixés, l'hypothèse d'extrapolation est souvent vérifiée.

Pour le ratio 4, comme les pourcentages croissent en avançant dans les étapes du protocole, l'hypothèse du BQ meilleur à *res1* qu'à *res2* est plus souvent vérifiée que l'hypothèse « les BQ à ces

deux résolutions sont assez proches ». Cette observation s'inverse en changeant de ratio. En effet, on observe pour le ratio 2 qu'on obtient un nombre d'images plus important qui vérifie cette hypothèse.

Pour la vérification de l'hypothèse d'extrapolation, l'influence de l'intervalle d'acceptabilité des distances est prépondérante. Nous avons choisi arbitrairement les bornes de ces intervalles par l'expérience. On remarque qu'un changement dans les bornes de cet intervalle aura des conséquences importantes sur les pourcentages finaux. Ainsi, si on élargit l'intervalle d'acceptabilité, on augmente d'autant plus le nombre d'images favorables.

Il reste maintenant à identifier les images défavorables afin de finir de cerner le domaine de validité de l'hypothèse d'extrapolation. Ensuite, nous explorons la possibilité de prédire les images qui seraient susceptibles de poser problème et celles qui systématiquement, ne vérifient l'hypothèse. Tout d'abord, nous proposons une analyse critique de notre approche empirique. Ensuite une série d'indices statistiques appliquée à des images en faveur ou non de l'hypothèse, nous a permis d'apporter les premières pierres dans une démarche de prédiction des images défavorables.

3.7 Critique de la démarche de l'étude empirique

3.7.1 Remarque sur la distance Q

Les tableaux correspondant au coefficient de corrélation et à l'indice Q présentent une forte similitude. Or, la distance Q est composée de trois termes, dont celui du coefficient de corrélation. Si les résultats sont redondants, cela signifie que les deux autres composantes de Q n'ont que très peu d'influence dans la situation dans laquelle nous nous sommes placés. La complexité de Q ne se justifie pas dans notre cas puisqu'un simple coefficient de corrélation suffisait. La raison de cette proximité est la suivante : hormis le coefficient de corrélation, Q est composé d'un terme basé sur les moyennes et un terme sur les écart-types des images MS de référence et fusionnées. Ces termes valent 1 si ces deux images possèdent la même moyenne et le même écart-type. Le modèle de fusion que nous avons utilisé conserve la moyenne des images d'origine, le terme relatif aux moyennes vaut donc 1. Seul l'autre terme, celui fonction des écart-types, pourrait différencier les deux indices. Cependant, le coefficient de corrélation est lui-même une fonction des écart-types, et la corrélation est ici très élevée puisque sa valeur est supérieure à 0.95. Plus la corrélation est élevée, plus les indices sont proches. En conclusion, lorsque le modèle de fusion ne transforme pas la moyenne des images, et lorsque le coefficient de corrélation est élevé (supérieur à 0.95), alors nous conseillons le recours à un simple coefficient de corrélation au lieu de l'indice Q . Aucun des autres tableaux de résultats ne présente une telle redondance.

3.7.2 Influence de l'algorithme de décomposition multiéchelle

Pour le ratio 2, on remarque que la vérification de l'hypothèse est très fréquemment moins contraignante en remontant dans les échelles. Cette constatation s'explique, au moins partiellement, par le choix de l'algorithme de décomposition multiéchelle. Il s'agit de la transformée en ondelettes 'à trous', qui a pour propriété d'être dyadique, non-orthogonale et non décimée. Elle a pour effet tout d'abord de produire des images qui ont toutes la même taille en termes de nombre de pixels, mais aussi de construire des plans de détails qui présentent des redondances.

Pour synthétiser des images qui servent de référence à des résolutions moindres, les images MS originales ont préalablement été suréchantillonnées pour avoir la même taille que l'image Pan, puis cet algorithme a été appliqué de manière récursive. Ainsi, même les images satellites de référence à *res1* résulte de l'application de cet algorithme. La redondance de l'information produite par l'utilisation de l'ondelette « à trous » se traduit par une corrélation non-nulle entre les détails. De même, les produits fusionnés synthétisés à partir de ces images ont tendance à présenter eux-mêmes des redondances, puisqu'ils sont créés à partir d'images déjà corrélées au départ.

On s'interroge maintenant sur le fait que les images de référence et produits fusionnés semblent être de plus en plus proches en remontant dans les échelles. La figure 3.4 illustre comment la transformée dyadique sépare le spectre fréquentiel des images en deux parties égales à chaque échelle.

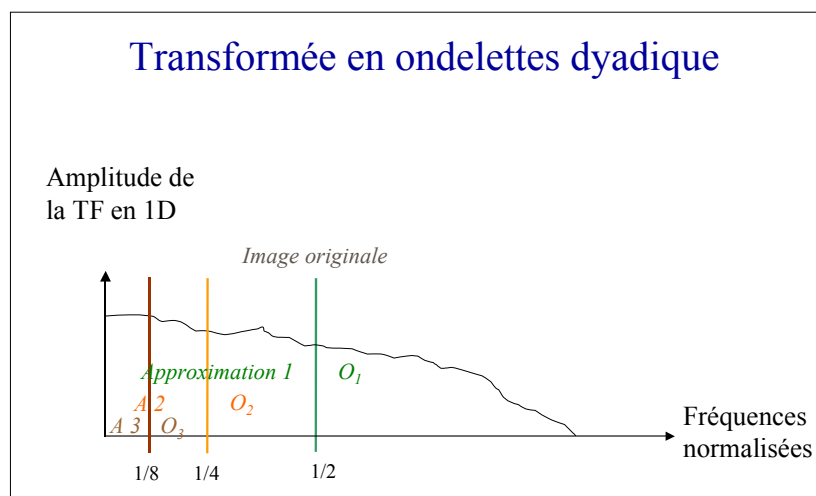


Figure 3.4 : spectre d'une image quelconque et résultats successifs de la transformée en ondelettes dyadique. Le spectre de Fourier représenté en fonction des fréquences normalisées est séparée en deux à chaque application de la transformée.

L'image originale (courbe en noire) crée par application d'une transformée en ondelettes, une image d'approximation à *res1* que nous appelons $A1$ pour 1^{ère} approximation, et résulte aussi en une image de coefficients d'ondelettes $O1$. Récursivement, les approximations $A2$ et $A3$ sont synthétisées de même que leurs plans d'ondelettes associés $O2$ et $O3$. Ainsi, pour une distribution fréquentielle uniforme dans l'image originale, la quantité d'information de $A1$ est moitié moindre que celle de l'image originale et deux fois plus importante que celle contenue dans $A2$. La transformée en ondelettes a pour propriété de conserver l'énergie à chaque échelle, ce qui se traduit mathématiquement par une variance constante des plans d'ondelettes avec les échelles. En effet, plus on remonte dans les échelles, moins il y a de coefficients d'ondelettes, mais plus ils sont par définition énergétiques, c'est-à-dire que leur valeur est élevée. Grâce à ces remarques, on comprend qu'un algorithme de fusion ait tendance à être moins performant à haute résolution spatiale qu'à des résolutions spatiales plus grossières.

Cette observation est en accord avec les résultats empiriques du ratio 2 où les BQ ont tendance à être de plus en plus proches en remontant dans les échelles. Or nous rappelons que l'hypothèse d'extrapolation concerne l'étude de la proximité entre un BQ à *res1* et ce même BQ à *res0*. Les résultats qui sont les plus proches de ceux que l'on aurait pu obtenir si on avait des références à *res0* sont ceux obtenus entre *res2-res1*, avec toutefois un point noir à souligner : les images à *res1* et *res2* ont toutes subi une étape de suréchantillonnage et plusieurs applications successives de la transformée en ondelettes. Elles ne correspondent pas à ce qu'aurait pu obtenir un capteur à ces deux résolutions spatiales.

3.7.3 Identification des images défavorables

Au vu de la séparation dyadique de l'information dans les échelles, nous avons étudié l'influence de la répartition fréquentielle des images. On s'attend à ce que si les hautes fréquences de l'image Pan sont très énergétiques, alors une grande quantité d'information est injectée au moment de la fusion. On suppose en s'appuyant sur notre expérience que la difficulté de synthèse est essentiellement liée à la distribution spectrale de l'énergie et donc à la complexité de la scène.

Il est aussi important d'introduire la notion d'échelle caractéristique des objets d'une scène. Puissant et Hirsch (2004) ont introduit la notion de Résolution Spatiale Optimale (RSO) qui consiste à considérer qu'il existe une échelle, ou un ensemble relativement étroit d'échelles, qui fournit la 'meilleure' représentation de chaque objet géographique observable. Par exemple, une image d'un

centre ville montre des objets différents (bâtiments, routes, immeubles, petites maisons) qui seront plus ou moins visibles ou identifiables en fonction de leur taille par rapport à la résolution spatiale d'acquisition. Si la résolution est supérieure à une dizaine de mètres, alors les bâtiments n'occupent que quelques pixels de l'image et il est difficile de distinguer les grands axes de circulation. Par contre, la ville comprend beaucoup d'objets qui possèdent une taille de quelques mètres, comme les voitures, les arbres des rues, la signalisation, les trottoirs, les fenêtres ... Il existe donc ce que l'on appelle une échelle caractéristique où ces objets vont commencer à être résolus de manière simultanée en augmentant la résolution spatiale. Nous nous attendons à ce qu'existe un saut en quantité d'information entre des images Pan à haute résolution spatiale par rapport à cette même modalité dégradée à la résolution de leurs modalités MS.

3.7.3.1 Distances utilisées pour étayer le discours

Pour l'étude de la complexité géométrique des images, nous avons retenu les distances suivantes :

- ♦ l'écart-type local moyen de l'image. Cet écart-type correspond à la moyenne des écart-types locaux ramenés à la moyenne locale. Par 'local', on entend un voisinage centré sur le pixel d'intérêt, fixé à une fenêtre de 5 par 5 pixels pour les résolutions des images auxquelles nous travaillons. La valeur moyenne finale est exprimée en pourcents ($\sigma_{RelLocal}(A0)$),
- ♦ le coefficient de corrélation entre les approximations à $res0$ et à $res1$ de la décomposition hiérarchique du Pan ($cc(A0, A1)$). Plus ces deux plans $A0$ et $A1$ sont éloignés, plus la valeur du $ccIntra$ sera faible,
- ♦ dans la même logique que pour le $ccIntra$, on calcule la différence de variances entre les approximations à $res0$ et $res1$ de l'image Pan, relative à la variance de l'approximation à $res0$ ($diffVarRel(A0, A1)$). La valeur est délivrée en %,
- ♦ enfin, l'entropie traduit la quantité d'information présente dans une image ; nous proposons de calculer la différence relative de l'entropie entre les approximations à $res0$ et à $res1$ du Pan ($diffHrel(A0, A1)$).

3.7.3.2 Images favorables et images défavorables

La figure 3.5 présente des extraits d'images Pan à leur résolution spatiale initiale. Les images b et d sont deux images qui ont largement validé l'hypothèse d'extrapolation et ce quel que soit le BQ utilisé. Au contraire, les images a et c représentent des zones exclusivement urbaines qui ont rejeté l'hypothèse. Les valeurs de plusieurs distances relevées à $res1$ étaient bien trop éloignées de celles obtenues à $res2$, et l'hypothèse a été invalidée pour ces deux images avec la plupart des BQ.

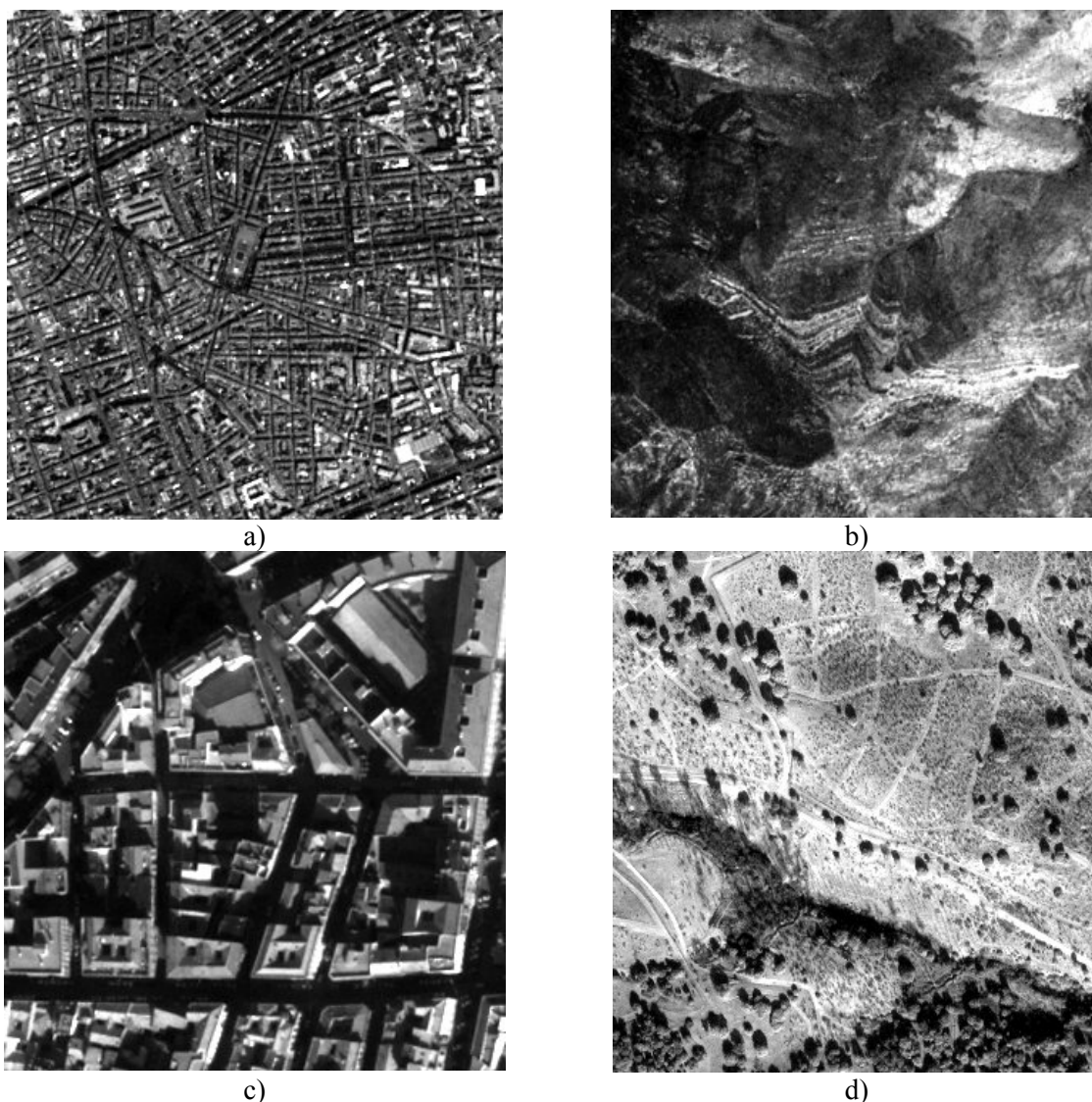


Figure 3.5 : modalités Pan. a) SPOT5, Marseille (France) centre ville, b) idem, collines provençales, c) Quickbird Madrid (Espagne) centre ville, d) idem, végétation éparse. Copyright CNES (2000) et Copyright Digital Globe (2000)

Nous avons calculé les distances retenues précédemment sur ces quatre extraits. Les résultats obtenus sont présentés dans le tableau 3.16.

$\sigma_{RelLocal}(A0) = 23 \%$ $cc(A0, A1) = 0.91$ $diffVarRel(A0, A1) = 46 \%$ $diffHrel(A0, A1) = 7 \%$	$\sigma_{RelLocal}(A0) = 9 \%$ $cc(A0, A1) = 0.98$ $diffVarRel(A0, A1) = 10 \%$ $diffHrel(A0, A1) = 1 \%$
a)	b)
$\sigma_{RelLocal}(A0) = 13 \%$ $cc(A0, A1) = 0.90$ $diffVarRel(A0, A1) = 41 \%$ $diffHrel(A0, A1) = 1 \%$	$\sigma_{RelLocal}(A0) = 6 \%$ $cc(A0, A1) = 0.95$ $diffVarRel(A0, A1) = 22 \%$ $diffHrel(A0, A1) = 2 \%$
c)	d)

Tableau 3.16 : résultats pour les quatre extraits de la figure 3.5.

On remarque que les différences de variance relative et les écart-types locaux sont beaucoup plus élevés pour les images présentant des centres villes que pour les autres. De même, la corrélation

entre les plans d'approximation à $res0$ et $res1$ est beaucoup plus élevée pour les images b et d. Par contre, la différence d'entropie relative n'offre pas de différence marquée entre les différentes images.

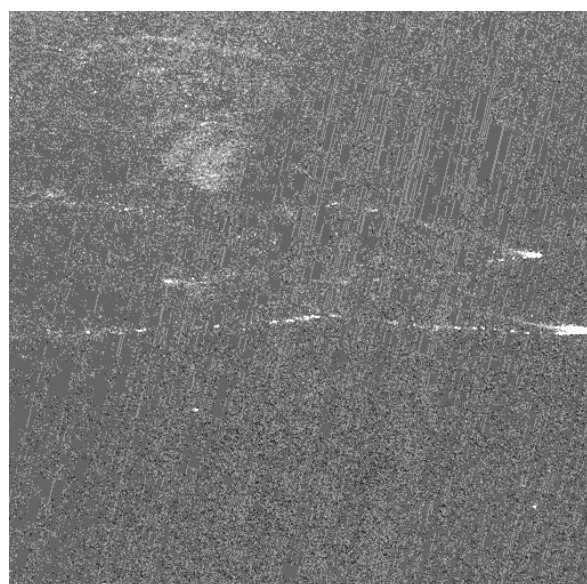
Nous n'avons pas appliqué ces statistiques à un nombre suffisant d'images pour pouvoir définir un seuil systématique de rejet ou d'acceptabilité d'une image donné vis-à-vis de l'hypothèse d'extrapolation. Par contre, on s'attend à ce qu'une Pan qui présente :

- ♦ un écart-type local élevé en moyenne,
- ♦ un coefficient de corrélation assez faible entre l'image Pan et son approximation à $res1$,
- ♦ et une différence en variance relative élevée entre l'image Pan et son approximation à $res1$,

nous indique que il y a une forte probabilité que l'hypothèse d'extrapolation ne soit pas vérifiée et donc un risque à admettre que le BQ de $res0$ est proche ou meilleur de celui à $res1$. Encore une fois, les conclusions sur l'hypothèse d'extrapolation dépendent des bornes que l'on s'est fixées pour affirmer que deux BQ sont « assez proches ».

3.7.3.3 Le cas particulier des images à fort signal sur bruit

La figure 3.6a montre un extrait de SPOT5 d'une zone située en pleine mer Méditerranée au large de Marseille. Les images de ce type rejettent aléatoirement l'hypothèse d'extrapolation. On observe des stries obliques inclinées d'environ 15° par rapport à la direction verticale, qui correspondent aux différences de gain des cellules CCD d'une même barrette. Ce bruit de capteur est un problème connu chez SPOT5 pour les capteurs de type push-broom, qui s'appelle le lignage. Ces stries donnent la direction de déplacement du satellite.



a)

$$\begin{aligned}\sigma_{RelLocal}(A0) &= 3 \% \\ cc(A0, A1) &= 0.54 \\ diffVarRel(A0, A1) &= 54 \% \\ diffHrel(A0, A1) &= 46 \%\end{aligned}$$

b)

Figure 3.6 : image SPOT5 - extrait de la mer Méditerranée, zone près de Marseille. Copyright CNES 2000.

Lorsque l'on visualise l'image entière de la ville de Marseille, ces stries sont invisibles. En accord avec nos remarques sur l'analyse visuelle, l'outil de visualisation a étalé la dynamique de l'image pour permettre de mieux distinguer les détails. La dynamique de cette image codée sur 8 bits est très faible, avec une moyenne de 20,3 et un écart-type de 0.79. Le rapport signal sur bruit est très élevé pour un tel extrait puisque les stries présentent des variations de plus ou moins un compte numérique. Mises à part les turbulences des sillons des deux bateaux sur la droite de l'image, les comptes numériques de l'image sont très faibles et regroupées autour de la moyenne de l'extrait.

Les valeurs des statistiques placées en b sur la même figure montrent un très faible écart-type local par rapport aux 4 extraits précédents. La corrélation entre les plans $A0$ et $A1$ est aussi beaucoup plus faible que pour les deux extraits correspondant aux deux zones urbaines problématiques. De même, la différence en variance relative et la différence relative en entropie dépassent les valeurs

trouvées pour les extraits problématiques. Donc on se retrouve avec une image qui présente une activité locale très faible, dont une grande partie est assignée à un bruit résiduel inhérent aux barrettes d'acquisition. Nous avons étudié de la même façon l'extrait de la rivière qui traverse la ville de Fredericton au Canada, acquis par Ikonos. Les mêmes tendances sont apparues : un faible écart-type local associé à un coefficient de corrélation faible et une différence en variance élevée.

Ces chiffres montrent qu'à cause du bruit important par rapport au signal d'intérêt, les approximations aux différentes échelles ne sont pas comparables les unes aux autres. Les BQ obtenus avec des produits fusionnés à différentes résolutions ont des valeurs aléatoires qui dépendent de la quantité de bruit et surtout de la répartition fréquentielle de ce bruit dans les échelles. De manière générale, lorsque l'on décroît la résolution, on diminue le niveau du bruit : dans le cas du lignage, le bruit est surtout situé dans les hautes fréquences spatiales, et disparaît donc en partie à cause du lissage utilisé pour synthétiser l'approximation à *res1*. Si une image ne présente que peu de structures, comme dans ces deux extraits, on conçoit que la qualité de la synthèse à basse résolution soit meilleure que celle dressée à haute résolution, ce qui invalide l'hypothèse. Il s'agit de cas particuliers : en présence de bruit et/ou lorsque le signal d'entrée est faible, on aura une plus grande probabilité d'invalider l'hypothèse d'extrapolation.

Notons que, de manière générale, ces images présentent un intérêt limité. De telles images peuvent être par exemple utilisées en détection de navires en pleine mer, mais une étape préalable de débruitage est nécessaire pour s'affranchir des bruits du capteur ou de compression.

3.7.4 Analyse complémentaire : vérification sur des extraits CNES

Les images à *res1* et *res2* sont plus proches en terme de contenu géométrique à cause de la transformée en ondelettes qui a permis de les synthétiser. En effet, nos références ont été synthétisées par l'analyse multiéchelle 'à trous' qui ne simulent pas la chaîne d'acquisition des satellites considérés. On se demande alors si nos résultats et nos conclusions sont généralisables au cas réel, à savoir si des références à *res1* et *res0* avaient été acquises directement par des capteurs satellites. Nous avons eu la possibilité de manipuler des images que nous a confiées le CNES, simulées dans le cadre du projet de lancement d'une constellation de satellites appelées Pléiades. Ces images ont été synthétisées à plusieurs résolutions spatiales à partir d'images aériennes acquises à 26 cm. La chaîne de simulation comprend la prise en compte de la *FTM* des capteurs, une transformée en ondelettes ainsi que plusieurs niveaux de compression. Nous avons procédé à une analyse complémentaire sur 2 extraits Pélican afin de comparer les résultats obtenus avec des références synthétisées par l'algorithme 'à trous' et en utilisant une transformée en ondelettes redondante et avec ceux obtenus par la chaîne de simulation du CNES qui donne des images plus proches de celles qu'un satellite est susceptible d'acquérir.

Nous avons à notre disposition des modalités MS à 40 cm, 80 cm, 160 cm et 320 cm. Il s'agit d'un cas idéal pour étudier l'évolution des BQ dans les échelles pour un ratio 2 ; les images MS à 320, 160 et 80 cm ont pu servir de modalités d'origine pour fusionner des images à 160, 80 et 40 cm, où les références étaient aussi disponibles. Nous avons synthétisé les images Pan à chaque résolution avec les coefficients de contribution spectrale relative de chaque modalité MS du satellite Quickbird, par la relation :

$$A = 1/(0.35+0.7+0.9+0.87) (0.87 B_{1,0}+0.9 B_{2,0}+0.7 B_{3,0}+0.35 B_{4,0}) \quad \text{Eq. 3.2}$$

où les indices des modalités MS 1, 2, 3 et 4 correspondent respectivement aux modalités PIR, rouge, verte et bleue.

Le tableau 3.17 suivant présente les résultats pour l'extrait 1 pour la vérification de l'hypothèse d'extrapolation. Les mêmes BQ que ceux de la vérification de l'hypothèse ont été utilisés, auxquels nous avons ajouté le BQ monomodal correspondant à la corrélation entre les hautes fréquences spatiales (*ccHF*) de la référence et de l'image fusionnée : c'est le BQ10. Ce tableau de résultats est composé de deux colonnes principales qui distinguent les comparaisons de BQ entre *res1* et *res0* et entre *res2* et *res1*. Ensuite, ces colonnes sont divisées en sous-colonnes, avec une sous-

colonne par modalité spectrale. A chaque case de résultat est assignée la valeur 1 ou 0. La valeur 1 indique que le BQ valide l'hypothèse d'extrapolation, et 0 l'invalidé.

	<i>res1-res0</i>				<i>res2-res1</i>			
	<i>B3</i>	<i>B2</i>	<i>B1</i>	<i>B0</i>	<i>B0</i>	<i>B1</i>	<i>B2</i>	<i>B3</i>
BQ1 = [<i>cc</i>]	1	1	1	1	1	1	1	1
BQ3 = [σ_{Rel} , <i>cc</i>]	0	0	0	1	1	1	1	1
BQ4 = [<i>diffvarRel</i> , σ_{Rel} , <i>cc</i>]	0	0	0	1	1	1	1	1
BQ5 = [<i>SAM</i>]	0				1			
BQ6 = [$M(V_{res})$, $\sigma(V_{res})$]	0				1			
BQ7 = [<i>ERGAS</i>]	0				1			
BQ8 = [<i>cc</i> , <i>SAM</i>]	0	0	0	0	1	1	1	1
BQ9 = intersection de toutes les distances	0	0	0	0	1	1	1	1
BQ10 = [<i>ccHF</i>]	1	1	1	1	1	1	1	1

Tableau 3.17 : résultats pour le premier extrait Pléiades. Pour un BQ donné, 1 signifie que l'hypothèse d'extrapolation est vérifiée, et sinon, la valeur de la case est 0.

Cet exemple contredit ce que l'on pourrait penser intuitivement, à savoir que la qualité devrait être meilleure en diminuant la résolution spatiale car le nombre de structures présentes dans l'image et leur complexité sont généralement supposés décroître lorsque la résolution spatiale diminue. En effet, les BQ montrent une qualité plus faible à la basse résolution spatiale *res2*. Cela confirme que la qualité ne dépend pas uniquement des échelles. Le problème est nettement plus complexe. Nous avons tenté d'identifier les raisons possibles pour lesquelles la qualité à *res1* est la meilleure, suivie de celle à *res0* pour finalement atteindre ses plus mauvais résultats à *res2*.

La qualité à *res1* est meilleure qu'à *res0*, ce qui signifie que cet extrait ne vérifie pas le cas 3. En effet, quelle que soit la distance, la qualité à *res1* est meilleure que celle obtenue à *res0*. Ainsi, les résultats vont fortement dépendre des bornes assignées à la définition de distances « assez proches ». On remarque que les BQ qui s'appuient sur les deux coefficients de corrélation valident systématiquement l'hypothèse. Ses bornes ne sont pas assez contraignantes, car le résultat est toujours très éloigné des bornes fixées. On note empiriquement qu'un seuil de +/- 0.015 serait suffisant pour le cas 2 de la vérification de l'hypothèse. Par contre, pour tous les autres BQ, l'hypothèse n'est pas valide.

Pour ce qui est de la comparaison entre *res2* et *res1*, la qualité à *res2* est moins bonne que celle à *res1* donc le cas 3 est vérifié pour toutes les distances. Le tableau 3.18 montre les valeurs obtenues pour la variance globale pour les images Pan à chaque résolution, et indique les différences de variance relative entre échelles successives. Comme les images Pan n'ont pas la même taille en terme de nombre de pixels, nous n'avons pu utiliser que des distances de la catégorie 1 correspondant à la différence entre des distances appliquées à une image. Nous n'avons donc pas pu utiliser le coefficient de corrélation.

<i>Approximation de l'image Pan</i>	<i>Variance globale</i>		
<i>A0</i>	57176	<i>DiffVarRel(A0, A1)</i>	20.03 %
<i>A1</i>	45722	<i>DiffVarRel(A1, A2)</i>	19.86 %
<i>A2</i>	36642	<i>DiffVarRel(A2, A3)</i>	30.20 %
<i>A3</i>	25576		

Tableau 3.18 : variances globales pour l'extrait 1 pour chaque approximation de l'image Pan. Les deux colonnes de droite correspondent à la différence en variance entre résolutions spatiales successives, relative à la variance de l'approximation à l'échelle plus fine.

La variance globale des images Pan diminue avec la résolution spatiale, donc la complexité de la scène diminue bien avec la résolution. Par contre, la différence en variance relative est plus élevée

entre les plans d'approximations $A2$ et $A3$ que pour les autres. Lors du passage de $A3$ à $A2$ on assiste à un saut de variances de plus de 10 % par rapport aux autres résolutions. Par contre, la valeur de la différence de variance entre $A0$ et $A1$ est similaire à celle entre $A1$ et $A2$ et le processus de fusion a rencontré plus de difficultés à haute résolution, ce qui est en accord avec ce que l'on pense intuitivement. Ainsi, la qualité de la fusion à une échelle donnée est probablement une fonction à la fois de l'échelle, mais aussi de la complexité de la scène MS à synthétiser à haute résolution spatiale.

Pour s'en convaincre, nous avons traité un second extrait CNES. Les résultats sont rapportés dans les tableaux 3.19 et 3.20.

	<i>res1-res0</i>				<i>res2-res1</i>			
	<i>B3</i>	<i>B2</i>	<i>B1</i>	<i>B0</i>	<i>B0</i>	<i>B1</i>	<i>B2</i>	<i>B3</i>
BQ1 = [<i>cc</i>]	1	1	1	1	1	1	1	1
BQ3 = [σ_{Rel} , <i>cc</i>]	1	0	0	1	1	1	1	1
BQ4 = [$diffvar_{Rel}$, σ_{Rel} , <i>cc</i>]	0	0	0	1	1	1	1	1
BQ5 = [<i>SAM</i>]	0				1			
BQ6 = [$M(V_{res})$, $\sigma(V_{res})$]	0				1			
BQ7 = [<i>ERGAS</i>]	0				1			
BQ8 = [<i>cc</i> , <i>SAM</i>]	0	0	0	0	1	1	1	1
BQ9 = intersection de toutes les distances	0	0	0	0	1	1	1	1
BQ10 = [<i>ccHF</i>]	1	1	1	1	1	1	1	1

Tableau 3.19 : résultats pour l'hypothèse d'extrapolation de l'extrait 2.

<i>Approximation de l'image Pan</i>	<i>Variance globale</i>		
<i>A0</i>	71074	$DiffVar_{Rel}(A0, A1)$	16.86 %
<i>A1</i>	59094	$DiffVar_{Rel}(A1, A2)$	15.48 %
<i>A2</i>	49949	$DiffVar_{Rel}(A2, A3)$	25.32 %
<i>A3</i>	37301		

Tableau 3.20 : variances globales et différences en variance entre résolutions spatiales successives, relatives à la variance de l'approximation à l'échelle plus fine, pour l'extrait 2.

Les résultats du tableau 3.19 sont parfaitement identiques à ceux de l'extrait 1 excepté pour la comparaison des BQ entre les résolutions $res1$ et $res0$, où pour le troisième BQ, la modalité PIR valide l'hypothèse. Sinon on retrouve une très bonne qualité de fusion à $res1$ générant une vérification systématique de l'hypothèse entre $res2$ et $res1$. La variance globale de l'image augmente bien avec la résolution spatiale. Par contre, on retrouve ce saut en différence de variance de 10 % en passant de ($A2$, $A3$) à ($A1$, $A2$). Malgré une complexité générale moindre, la fusion a donc tendance à être moins performante à $res2$ qu'à $res1$ puisque la quantité d'information qui sépare les approximations successives est plus importante entre $A3$ et $A2$. Les pourcentages de la différence en variance relative des plans ($A0$, $A1$) et ($A1$, $A2$) sont très proches. Ainsi, comme la scène est plus complexe à haute résolution spatiale, les BQ à $res0$ sont plus mauvais qu'à $res1$.

Plusieurs conclusions peuvent être tirées :

- ♦ en aucun cas une étude entre BQ à $res2$ et $res1$ nous permet de conclure à un quelconque comportement entre $res1$ et $res0$,
- ♦ la performance d'une méthode de fusion ne dépend pas uniquement de l'échelle considérée,
- ♦ la complexité de la scène à synthétiser joue un rôle dans la qualité du produit fusionné final. Des paramètres calculés à partir de l'image Pan et de son approximation à la résolution spatiale d'origine des images MS sont susceptibles de nous renseigner sur la difficulté de fusion. Cependant, nous n'avons pas été en mesure de définir une distance et un critère associé permettant de prédire la qualité de la fusion.

3.8 Conclusion et perspectives

Plusieurs interrogations ont été soulevées, à savoir si le mode opératoire choisi permettait de généraliser au cas réel. En effet, une seule méthode de fusion a été testée, et il serait intéressant de voir si le comportement observé avec la méthode *ATWT-M2* est généralisable aux autres méthodes de fusion. De plus, le résultat est très dépendant du BQ considéré, et on a tendance à augmenter la probabilité d'invalidier l'hypothèse d'extrapolation en augmentant le nombre de distances qui composent le BQ. Nous pouvons désormais apporter des éléments de réponse aux trois questions posées en introduction, à savoir :

- ♦ Est-ce que cette hypothèse est souvent vérifiée ? Avec les nouveaux critères pour les distances, la campagne de tests menée nous permet de conclure qu'il est raisonnable de supposer que pour la plupart des images, la qualité des images synthétiques à *res0* est assez proche et/ou moins bonne de celle trouvée à *res1*, et donc l'hypothèse d'extrapolation est vérifiée. Notons toutefois que ces conclusions ne sont valables que dans le cas de la méthode de fusion testée.
- ♦ Quel est son domaine de validité ? Les images qui ne valident pas l'hypothèse sont ceux où la quantité d'information à injecter de l'image Pan vers l'image MS est la plus importante. De manière générale, les scènes complexes rencontrent plus de difficultés pour valider l'hypothèse que des scènes plus simples en terme de quantité de signatures spectrales et structures géométriques. Il est aussi important de souligner que le domaine de validité de l'hypothèse dépend du BQ considéré.
- ♦ Est-il possible de prédire, à coup sûr, les images pour lesquelles l'hypothèse est ou non, valide ? A ce stade de l'étude, il n'est pas possible de prédire à coup sûr les images pour lesquelles l'hypothèse serait valide. La qualité d'un produit de fusion est une fonction qui dépend de l'échelle, mais aussi de la quantité d'information à injecter dans les modalités MS. Pour un BQ et une méthode de fusion donnée, des travaux supplémentaires sont nécessaires pour être susceptible de prédire l'acceptabilité ou non pour une image donnée.

Les analyses nous ont aussi permis de proposer de nouveaux critères, moins restrictifs pour certaines distances, et plus contraignants pour d'autres. Le tableau 3.21 propose de nouveaux intervalles aux distances.

<i>Distances</i>	<i>Intervalle de l'hypothèse « assez proches »</i>	<i>Rappel des valeurs idéales</i>
<i>diffVarRel</i>	+/- 15 %	0
<i>σ_{Rel}</i>	+/- 10 %	0
<i>cc</i>	+/- 0.015	1
<i>ccHF</i>	+/- 0.015	1
<i>ERGAS</i>	+/- 2	0
<i>SAM</i>	+/- 2°	0

Tableau 3.21 : nouvelles bornes d'acceptabilité et rappel des valeurs idéales pour chaque distance.

Les bornes des intervalles des deux coefficients de corrélation étaient trop permissives ; ils avaient tendance à accepter toutes les images, mêmes pour les scènes les plus complexes. Les contraintes sur les autres distances ont été relâchées.

Ces résultats nous ont aussi montré les limites de certaines distances que nous avons proposées dans le chapitre précédent. En effet, les statistiques concernant la différence des normes des vecteurs spectraux et celles de la norme du vecteur résultant ne sont pas de bonnes distances de la manière dont elles ont été définies. En effet, d'après Wald (2002), une distance multimodale doit être indépendante du nombre de modalités MS, de la dynamique des images ainsi que de la résolution

spatiale. Or pour le moment, ces distances sont sensibles à ces divers aspects, ce qui rend impossible la définition d'un critère cohérent. Nous avons proposé :

$$diffNorms(i, j) = ||B(i, j)|| - ||B^*(i, j)|| \quad \text{Eq. 3.3}$$

$$V_{res}(i, j) = ||B(i, j) - B^*(i, j)|| \quad \text{Eq. 3.4}$$

Pour devenir un indice fiable, les composantes des vecteurs multispectraux doivent être réécrites pour prendre en compte la dynamique des images et le nombre de modalités. Si σ_k est l'écart-type de la $k^{ième}$ modalité MS et N le nombre total de modalités, alors :

$$||B(i, j)|| = \sqrt[3]{1/N \sum (B_k(i, j)/\sigma_k)^2} \quad \text{Eq. 3.5}$$

Chaque coordonnée du vecteur MS est normalisée par rapport à l'écart-type de la modalité MS considérée. Ainsi, si une modalité possède une dynamique deux fois plus importante qu'une autre, elle influence la distance dans les mêmes proportions que les autres. De plus, une division par rapport au nombre de modalités MS évite un accroissement de la valeur de la distance avec une augmentation du nombre de modalités. De plus, concernant la différence des normes, nous proposons de considérer la valeur absolue de cette grandeur pour éviter les valeurs négatives. De plus amples recherches doivent être menées sur ces distances.

IV Un nouvel outil pour l'évaluation de la qualité géométrique, basé sur la FTM

Ce chapitre est consacré à la proposition d'une nouvelle distance pour l'évaluation de la qualité géométrique des produits de fusion, basée sur l'estimation de la fonction de transfert de modulation des capteurs. Cette distance doit être utilisée dans les mêmes conditions que les autres distances définies au chapitre 2, à savoir qu'elle doit être appliquée pour mesurer la proximité entre une image fusionnée et son homologue de référence. Nous rappelons que nous ne recommandons pas la démarche usuellement rencontrée dans la littérature qui consiste à appliquer les distances géométriques entre la modalité fusionnée et la modalité Pan.

4.1 Introduction

4.1.1 Etalonnage et fonction de transfert de modulation

L'objectif de la fusion d'images est de simuler ce qu'un capteur multispectral pourrait observer s'il possédait la haute résolution spatiale. Les caractéristiques du capteur jouent un rôle prépondérant lors de la simulation d'images aux différentes échelles, et notamment l'étalonnage radiométrique et la fonction de transfert de modulation.

Pour rendre les grandeurs comparables, il faut transformer les comptes numériques en grandeurs physiques, ici en luminance. En effet, le capteur transforme une énergie lumineuse incidente en niveaux de gris selon un procédé qui va dépendre de la qualité de la barrette d'acquisition, de la quantification... Lorsque des formules mathématiques concernent plusieurs capteurs, l'étalonnage est un alignement radiométrique qui permet de quantifier la réelle contribution de chacun d'entre eux en l'associant à une grandeur physique comparable. L'étalonnage est souvent une fonction linéaire :

$$L = \alpha CN + \beta \quad \text{Eq. 4.1}$$

où L est la valeur en luminance, CN correspond aux comptes numériques et α et β sont deux coefficients d'étalonnage donnés pour chaque image par le fournisseur. Certaines méthodes de fusion n'utilisent pas l'étalonnage et les résultats obtenus montrent souvent une très forte distorsion spectrale (Wald 2000). Les plus évoluées, comme *RWM* (Ranchin *et al.* 1993), prennent en compte cet étalonnage.

Dans ce chapitre, nous avons choisi de nous intéresser plus particulièrement à la seconde caractéristique importante des capteurs qu'est la fonction de transfert de modulation (*FTM*).

4.1.2 Définition de la FTM

4.1.2.1 Caractérisation de la résolution spatiale d'un capteur

La résolution spatiale d'un capteur est une caractéristique importante qui permet l'analyse de sa performance. Il s'agit d'une grandeur complexe qui a fait l'objet de nombreuses tentatives de définition. La résolution spatiale est une grandeur qui caractérise la capacité d'un système imageur à décrire les détails les plus fins qu'il est capable de détecter. De manière générale, sa valeur est donnée par la taille d'un pixel de l'image projetée au sol. La qualité d'une image est très souvent caractérisée par des critères radiométriques usuels tels que la moyenne, le contraste, la luminosité, la variance du bruit, ou la résolution radiométrique. Seuls quelques outils rendent compte de la performance

géométrique des capteurs. Les plus fréquents sont ceux qui s'appuient sur la réponse d'un système optique à une impulsion lumineuse (*Point Spread Function*) (Li 2000, Lei et Tiziani 1989), le gradient ou d'autres filtres passe-haut (Zhou 1998), le *Ground Sample Distance* (*GSD*, ou plutôt *Equivalent GSD*) (Vrabel 2000), la *FWHM* (*Full Width at Half Maximum*), grâce à des analyses visuelles (IRARS 1996) ou encore par le rapport signal sur bruit (*SNR*, Dial *et al.* 2003).

Utilisés indépendamment les uns des autres, ces outils sont incomplets car ils n'extraient qu'une partie de la performance, et négligent ainsi les informations relatives aux autres fréquences spatiales transmises par le capteur. La *FTM* permet la caractérisation de la qualité géométrique globale des capteurs, en offrant une mesure précise effectuée dans le domaine des fréquences.

4.1.2.2 Définition de la *FTM* et relation avec la synthèse d'images MS à haute résolution spatiale

La *FTM* traduit comment le capteur atténue le contraste de l'image en fonction des fréquences spatiales. L'atténuation est causée par la taille finie des détecteurs, les mouvements du capteur, la diffraction, les aberrations optiques, la dispersion de l'atmosphère turbulente et les parasites liés à l'électronique (Holst 1995). Seule la moyenne de l'image est parfaitement préservée, et la *FTM* tend à réunir tous les comptes numériques autour de cette moyenne. De plus, plus le bruit *B* de la chaîne d'acquisition est important, plus la *FTM* aura tendance à présenter un écart-type faible. Si la fréquence spatiale d'un objet est supérieure à la fréquence de coupure du capteur, alors son image sera représentée par un éclaircissement uniforme.

Par définition, la *FTM* d'un système est la Transformée de Fourier (*TF*) de sa réponse impulsionnelle ou *PSF*. Cette réponse consiste en une tache très lumineuse en son centre et devenant de plus en plus sombre en s'écartant du centre :

$$FTM = TF(PSF) \quad \text{Eq. 4.2}$$

Une image, par exemple la modalité B_{kl} , correspond à la convolution entre la scène et la réponse impulsionnelle (PSF_{resl}) du capteur. L'indice *resl* de la *PSF* signifie que nous nous plaçons à la résolution *resl* d'acquisition des modalités MS :

$$B_{kl} = PSF_{resl} * \text{scène} \quad \text{Eq. 4.3}$$

où $*$ est l'opérateur de convolution. Dans le domaine de Fourier, la convolution se transforme en produit simple (théorème de Parseval). Cette expression devient :

$$TF(B_{kl}) = TF(PSF_{resl}) TF(\text{scène}) \quad \text{Eq. 4.4}$$

C'est-à-dire,

$$TF(B_{kl}) = FTM_{resl} TF(\text{scène}) \quad \text{Eq. 4.5}$$

Ainsi, la transformée de Fourier de l'image B_{kl} s'obtient en multipliant la *FTM* de ce capteur (donc à *resl*) par la *TF* de la scène.

Si on définit par PSF_{res0} et FTM_{res0} respectivement la réponse impulsionnelle et la *FTM* de la $k^{\text{ième}}$ modalité MS B_{k0}^* synthétisée à *res0*, alors :

$$B_{k0} = PSF_{res0} * \text{scène} \quad \text{Eq. 4.6}$$

et,

$$TF(B_{k0}) = FTM_{res0} TF(\text{scène}) \quad \text{Eq. 4.7}$$

Pour faciliter l'écriture, nous nous plaçons dans le cas non discret. Le rapport entre la transformée de Fourier de la modalité à *res0* et la modalité à *resl* :

$$TF(B_{k0})/TF(B_{kl}) = FTM_{res0}/FTM_{resl} \quad \text{Eq. 4.8}$$

La synthèse d'images MS à haute résolution spatiale peut se ramener alors à un problème de déconvolution. Il s'agit d'un problème mal posé : si on est capable de connaître le terme FTM_{res0}/FTM_{resl} , alors on sera en mesure d'estimer B_{k0}^* à partir de B_{kl} .

$$TF(B_{k0}^*) = TF(B_{kl}) FTM_{res0}/FTM_{resl} \quad \text{Eq. 4.9}$$

4.1.2.3 Modélisation de la FTM

Le modèle de *FTM* que nous avons adopté résulte de l'influence de trois éléments : l'optique, l'échantillonnage et l'effet de filé. Dans cette modélisation, nous ne tenons pas compte par exemple de l'influence de l'atmosphère et du bruit d'acquisition. Notons qu'en présence d'un bruit additif, cette expression de la *FTM* du capteur d'acquisition de la modalité B_{kl} devient :

$$TF(B_{kl}) = FTM_{resl} TF(scène) + B(f) \quad \text{Eq 4.10}$$

où $B(f)$ est la transformée de Fourier du bruit d'acquisition.

La convolution de ces trois termes aboutit à la fonction finale *FTM* finale :

$$FTM = FTM_{optique} * FTM_{échantillonnage} * FTM_{filé} \quad \text{Eq 4.11}$$

Le premier terme modélise l'optique de la chaîne d'acquisition. Cette modélisation repose sur la théorie de la diffraction : l'image d'un objet ponctuel à travers un objectif (ou pupille) parfaitement circulaire de diamètre D_p , sans défaut ni obstruction, est une tache circulaire entourée d'anneaux concentriques. Cette image est appelée tache d'Airy (figure 4.1) ou tache de diffraction car elle correspond au plus petit élément d'image perceptible à travers l'instrument optique.

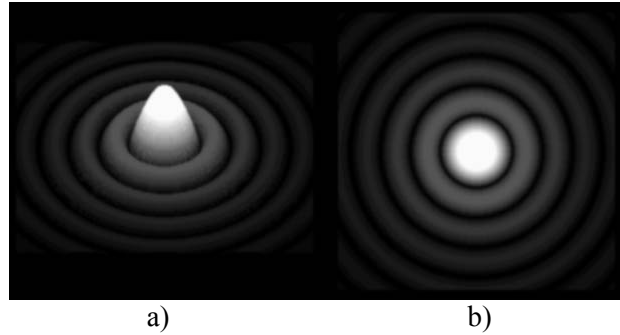


Figure 4.1 : tache d'Airy. a) vue avec une élévation de 45°, b) vue de dessus.

La figure 4.1a représente la tache d'Airy en trois dimensions vue avec un angle d'élévation de 45°. A droite en b, cette tache figure vue de dessus. On peut observer les cercles concentriques, centrés sur la valeur maximale en blanc sur l'image. En s'écartant du centre, les cercles ont une radiométrie de plus en plus atténuée. Si λ correspond à la longueur d'onde du rayonnement incident, la fréquence de coupure optique $F_{Coptique}$ s'écrit comme le ratio suivant :

$$F_{Coptique} = D_p / \lambda \quad \text{Eq. 4.12}$$

Au delà de cette fréquence, les fréquences spatiales ne sont plus transmises. La $FTM_{optique}$ est donnée par :

$$FTM_{optique}(f, \lambda) = \exp(-\lambda |f|^\alpha) \quad \text{Eq 4.13}$$

où α est plus grand que 1 et correspond à un facteur d'atténuation. Notons que les autres aberrations possibles en optique, telles que l'influence de l'atmosphère, la turbulence, ou les déformations ne sont pas prises en compte dans ce modèle.

Le deuxième terme correspond à la modélisation de l'échantillonnage. D'après le théorème de Shannon, la transformée de Fourier d'un signal numérique $E(f)$ est limitée par la fréquence d'échantillonnage qui a pour effet de périodiser son spectre en fréquences. Si ce signal comporte des fréquences spatiales supérieures à la demi-fréquence d'échantillonnage appelée fréquence de Nyquist (f_N), alors on obtient une surévaluation des fréquences : c'est le phénomène du repliement spectral, connu aussi sous le nom d'*aliasing*. La figure 4.2 permet de localiser les fréquences de Nyquist et d'échantillonnage sur la courbe qui donne la transformée de Fourier d'un signal d'entrée $E(f)$ en fonction des fréquences spatiales.

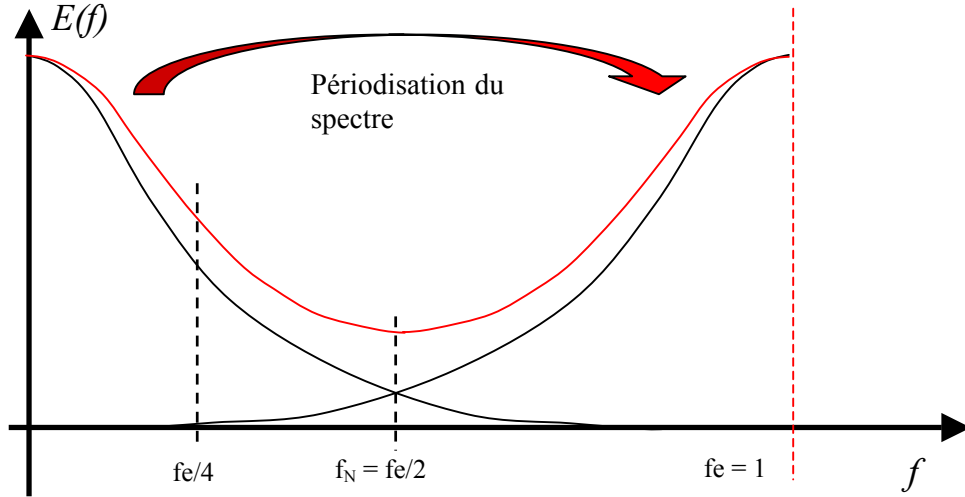


Figure 4.2 : TF d'un signal d'entrée $E(f)$ échantillonné (courbe rouge). La courbe rouge témoigne du repliement spectral qui correspond à la somme des courbes noires, signal de l'image de départ qui a été périodisé par l'échantillonnage tous les $(k f_e)$ avec $k \in \mathbb{Z}$.

Soit D_d la taille des détecteurs de la barrette d'acquisition. Un échantillonnage contigu peut être vu comme un échantillonnage parfait du signal filtré par un filtre moyenneur. Si on néglige les différences de sensibilité des CCD, on peut modéliser ce filtrage par une fonction sinus cardinal :

$$FTM_{\text{échantillonnage}}(f) = \text{sinc}(D_d f) \quad \text{Eq 4.14}$$

Le troisième terme est l'effet de filé. Il s'agit d'un effet produit par le défilement du satellite. En effet, une cellule d'une barrette CCD récolte les photons émis par une scène en un temps donné, que l'on appelle le temps d'intégration t_{int} . Pendant ce temps, le satellite se déplace, et l'effet de filé se comporte comme un filtre intégrateur dans la direction de défilement. Une asymétrie dans la représentation de la PSF est produite, donnant des estimations de la FTM différentes selon la direction considérée. En effet, la résolution a tendance à être plus faible dans la direction de défilement comparée à sa direction orthogonale (Latry 2003). Si D_t correspond à la distance parcourue dans le plan image et que v_{sat} est la vitesse du satellite, alors :

$$Dt = v_{sat} t_{int} \quad \text{Eq. 4.15}$$

$$FTM_{\text{filé}}(f) = \text{sinc}(D_t f) \quad \text{Eq. 4.16}$$

4.1.2.4 Réticence des constructeurs de capteurs spatiaux

Dans le domaine de l'observation de la Terre, la FTM est une information très difficile à obtenir avec précision. Les constructeurs et fournisseurs d'images sont réticents pour fournir la forme de la courbe et les fréquences caractéristiques associées. La raison est que la FTM est directement liée à la qualité intrinsèque des produits vendus. Par exemple, la documentation fournie par SPOT Image (Anonymous 1986) reste très vague au sujet de la FTM . Elle donne des spécifications pour la fréquence de coupure du canal Pan à 10 m de résolution sans pour autant donner d'information sur la façon dont elle a été calculée. De plus, aucune information n'a été trouvée à propos des capteurs MS. Un autre exemple concerne les spécifications de Quickbird : les seules informations disponibles sont « high contrast FTM » et « 61-centimeters GSD for Panchromatic and 2.44-meters GSD for multispectral » (<http://www.hgiis.com/documents/en/Product%20Guide%20v42Asia%2005-20-2004.pdf>). La FTM n'est jamais décrite avec précision par les constructeurs. De plus, l'usage de nombreux termes tels que pouvoir de résolution, fréquence de coupure, valeur de la FTM minimale autorisée à la fréquence de Nyquist, ou encore pourcentage d'atténuation du contraste initial à cette même fréquence (Dial *et al.* 2003), crée une confusion parmi les lecteurs et les utilisateurs.

Nous avons donc développé notre propre méthode d'estimation de la FTM . Connaître la FTM , c'est mieux connaître le capteur et la qualité des images qu'il fournit. L'importance de prendre en compte la FTM dans la fusion et les modèles de décomposition multiéchelles a déjà été soulignée à

plusieurs reprises (Aiazzi *et al.* 2002). En étudiant cet outil, nous avons eu l'idée de l'utiliser afin de décrire la qualité géométrique d'une image, et plus particulièrement de proposer une distance pour la qualité géométrique des produits de fusion.

4.2 Présentation de notre méthode d'estimation de la FTM

4.2.1 Limite des méthodes existantes

L'estimation de la *PSF* consiste à dessiner un point infiniment localisé et d'amplitude infinie et de voir comment ce point est transmis par le système d'acquisition. En pratique, on peut envisager de dessiner un point noir sur fond blanc pour un scanner ou un appareil photo numérique. Le problème devient nettement plus complexe pour un capteur satellite. L'idéal est d'avoir la possibilité de positionner des spots lumineux dans des scènes à radiométrie à peu près homogène telles que des champs, comme l'ont entrepris Léger *et al.* (2002). Mais l'utilisation de cibles artificielles limite le champ d'application. De plus, le spot lumineux servant à l'estimation de la *PSF* peut être noyé dans le bruit d'acquisition et sa réponse dépend totalement de la localisation du pic lumineux sur la grille d'échantillonnage.

Les deux problèmes fondamentaux dans l'estimation de la *FTM* sont le bruit et l'aliasing. Le bruit a pour effet d'atténuer la *PSF* ce qui sous-évalue la *FTM*. Au contraire, le repliement spectral engendre une surévaluation des fréquences, plus particulièrement aux environs de la fréquence de Nyquist. Pour lutter contre ces deux problèmes, Ryan *et al.* (2003) ont eu recours à un modèle paramétrique de *PSF*. D'autres possibilités ont été proposées pour augmenter le nombre de valeurs pour décrire la *PSF* ; soit en disposant de plusieurs images du même spot lumineux légèrement décalées par rapport à la grille d'échantillonnage, soit d'une unique image mais avec plusieurs spots aléatoirement disposés dans la scène. Ainsi, on pourrait obtenir une version suréchantillonnée et moyennée de la *PSF* à partir de laquelle la *FTM* pourrait être reconstruite sans effet indésirable d'aliasing et ainsi limiter le bruit. Cependant, encore une fois cette démarche est difficilement concevable avec l'imagerie satellite.

La méthode d'estimation de la *FTM* que nous proposons n'est pas nouvelle en soi. Elle correspond à une série d'étapes déjà existantes dans la littérature, par exemple dans les travaux de Léger *et al.* (2004). L'innovation réside dans le choix, que nous espérons optimal, de chacune des étapes provenant de méthodes de lecture différentes. La validation s'appuie sur les études menées sur l'imagerie Ikonos par Ryan *et al.* (2003) et Dial *et al.* (2002).

4.2.2 Principe de l'estimation de la FTM

Une manière de s'affranchir de l'estimation de la *PSF* est de la simuler à partir des informations déjà disponibles dans l'image. La méthode habituelle est de calculer la *PSF* en utilisant une approche basée contours. En effet, si on extrait une ligne perpendiculairement à un contour d'une image, on obtient un profil qui correspond à l'intégrale de la *PSF*. La figure 4.3 montre un profil centré sur l'indice 0 avec en ordonnées les luminances ou niveaux de gris en valeurs normalisées.

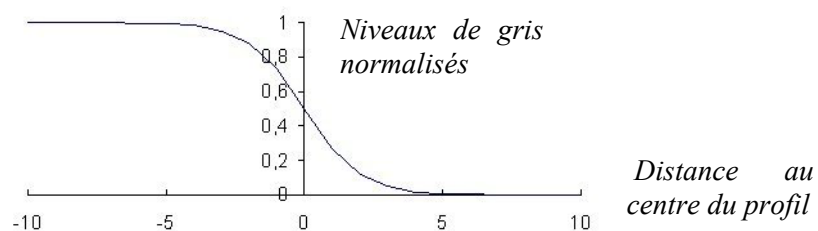


Figure 4.3 : profil d'une transition.

Il suffit alors de dériver ce profil pour obtenir la *PSF*. En théorie, ce contour doit correspondre à une transition idéale, c'est-à-dire pouvoir couvrir la dynamique de l'image (Ryan *et al.* 2003, Anuta *et al.* 1984). D'après le même principe que pour le spot lumineux, l'image par un système

d'acquisition d'un contour parfait est une transition qui n'est pas parfaite. Elle apparaît beaucoup plus progressive et sa progression donne des informations sur la manière dont le système traite les contours. Ceci est illustré par la figure 4.4. A gauche (a), un contour parfait. A droite, le même contour vu par le système d'acquisition. Le contour très net est transformé en un dégradé.

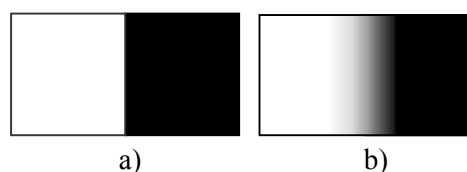


Figure 4.4 : a) contour parfait, b) contour perçu à travers un système d'acquisition.

La dérivation d'une section de cette image prise perpendiculairement à ce contour aboutit à la *PSF*. Enfin, la Transformée de Fourier de la *PSF* permet d'obtenir la *FTM* dans une direction donnée.

Cette approche basée contour permet efficacement de lutter contre le problème du bruit puisqu'elle a l'avantage de pouvoir fournir une version du profil par ligne de l'image ; les profils seront ensuite moyennés afin de s'affranchir du bruit. Cependant, le problème de l'aliasing demeure. On souhaite, comme dans le cas du spot lumineux, pouvoir bénéficier de versions de ce contour légèrement décalées par rapport à la grille d'échantillonnage pour chaque ligne. Pour cela, il suffit d'utiliser un contour non plus aligné avec l'un des bords de l'image mais oblique comme le montre la figure 4.5.

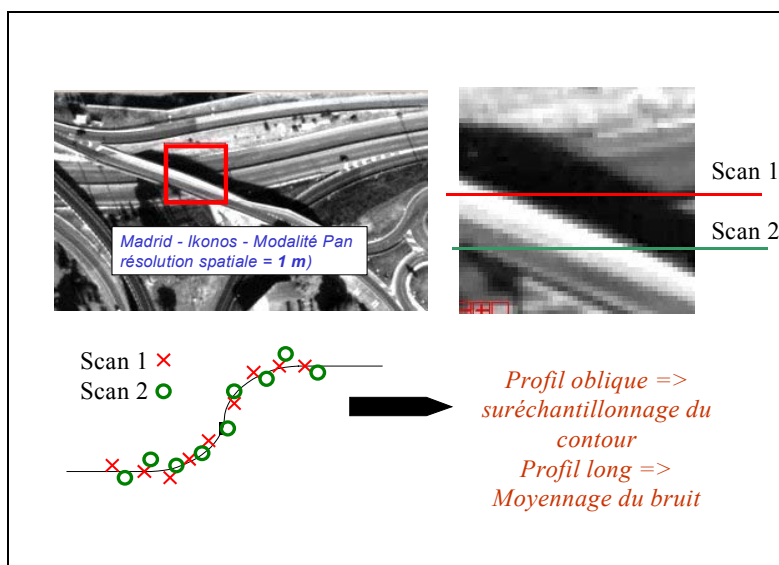


Figure 4.5 : principe de l'approche basée contour.

Cette figure illustre comment un contour qui présente une transition oblique rectiligne et avec un bon gradient de radiométrie permet de créer un suréchantillonnage du profil. La transition d'intérêt est formée d'une route de radiométrie élevée et de son ombre de radiométrie plus faible et relativement homogène. Chaque ligne présente un échantillonnage différent du profil comme le montrent les scans 1 et 2. Certains échantillonnages seront similaires et leurs profils vont se superposer, alors que d'autres seront décalés et permettront de décrire plus finement le contour.

Cette approche avait déjà été proposée par Ryan *et al.* (2003) et Dial *et al.* (2003) mais ils utilisaient des cibles artificielles manuellement disposées sur la trajectoire du satellite Ikonos. Notre objectif est de proposer une méthode simple et efficace d'estimation, utilisable par tous avec les images qui sont à notre disposition. Notre alternative se base donc sur la lecture des contours naturels d'une image et permet simultanément un suréchantillonnage du profil avec un moyennage des échantillons qui se superposent. Les différentes étapes d'estimation de la *FTM* sont décrites ci-après. Chacune de ces étapes fait appel à des outils simples et connus dans le domaine du traitement d'images. La méthode d'estimation est donc facilement reproductible.

4.2.3 Calcul de l'équation de la droite du contour

La méthode débute par une sélection manuelle de l'utilisateur d'un contour d'intérêt. Le contour doit être long et rectiligne et posséder une faible inclinaison par rapport à la verticale, de l'ordre d'une dizaine de degrés. La transition doit être de telle sorte que son minimum et son maximum se rapprochent de la dynamique de l'image. Ces contraintes restreignent considérablement le jeu d'images utilisables pour un capteur donné. Peu d'images ont pu être retenues et a fortiori peu d'extraits. Plus l'image est bruitée, plus il est conseillé de favoriser des contours allongés. Nous travaillons sur des images possédant une haute résolution spatiale de l'ordre du mètre. Ces contours pourront être des arêtes de hangars, des ponts, les contours d'un stade, les bords d'une route ou encore d'une piste d'aviation. Les objets sélectionnés dépendent fortement de la résolution spatiale de l'image puisqu'il est nécessaire d'avoir un contour d'au moins quelques pixels de large. Généralement, les ombres des bâtiments procurent de bonnes différences d'amplitude et une homogénéité satisfaisante. Ainsi pour des résolutions de quelques mètres où ils sont susceptibles d'occuper plus de quelques pixels de l'image, s'ils présentent aussi une orientation satisfaisante, ces objets peuvent constituer de bons contours potentiels pour l'estimation de la *FTM*.

La figure 4.6a représente un exemple de contour que nous avons sélectionné dans une image Ikonos de la ville de Fredericton (Canada). Pour obtenir l'équation de la droite représentant le contour, on doit dériver le contour, par exemple par un filtre de Sobel. Néanmoins, comme la dérivation revient à une convolution par un filtre passe-haut, il faut se prémunir du bruit pour s'affranchir des structures parasites en appliquant un filtrage passe-bas horizontal préalable (figure 4.6b). Le résultat obtenu est placé en figure 4.6c. Si on ne prend pas garde à diminuer le bruit, on obtient les artefacts de la figure 4.6d.

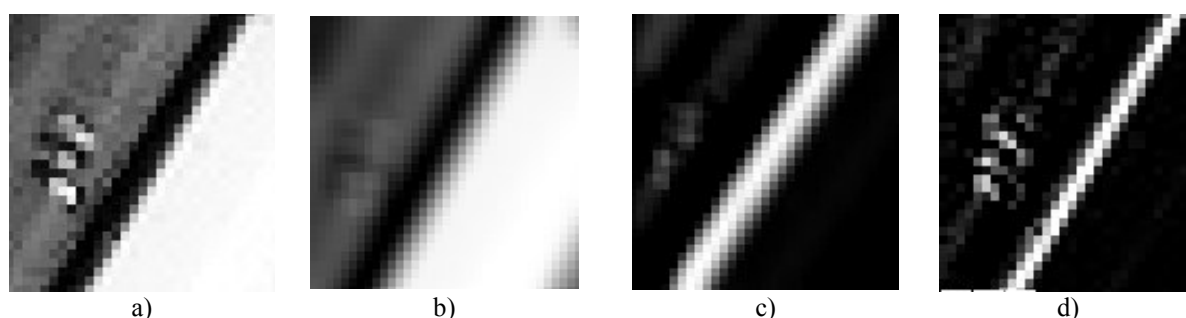


Figure 4.6 : image Ikonos – arête d'un bâtiment de Fredericton (Canada). a) contour extrait manuellement, b) moyennage du contour par un filtre passe-bas, c) dérivation numérique par un opérateur de Sobel à une dimension, d) résultat de la dérivation sans filtrage passe-bas préalable. Copyright Space Imaging 2002.

Ensuite, on utilise la transformée de Hough pour passer du domaine spatial à celui des paramètres des fonctions que l'on cherche à identifier dans une image. Ici, on cherche à détecter une droite. Pour ce type d'objets, à chaque couple (x, y) de l'image (figure 4.7a) correspond une sinusoïde dans l'espace de Hough (figure 4.7b). En effet, cette transformée utilise une représentation paramétrique en coordonnées polaires des courbes et objets à détecter.

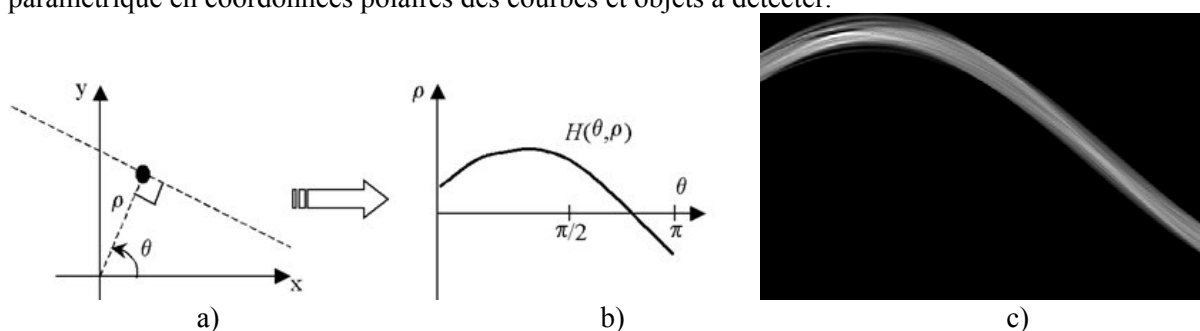


Figure 4.7 : principe de la transformée de Hough, a) du domaine de l'image, vers b) et c) celui de Hough.

L'équation d'une droite s'écrit $y = ax + b$ dans le domaine spatial et devient $y = \rho \sin\theta$ et $x = \rho \cos\theta$ en coordonnées polaires. On applique donc cette transformée à l'image de la figure 4.7c. Ensuite, on applique un seuil aux coefficients de Hough pour sélectionner les plus élevés, c'est-à-dire aux endroits où le plus de sinusoides s'intersectent (figure 4.8c) ; un critère de sélection possible des coefficients de Hough est de choisir les coefficients de Hough qui ont une valeur supérieure au percentile 85 % de cette image.

La prochaine étape consiste donc à appliquer la transformée inverse pour revenir dans l'espace de l'image (figure 4.8a). Un seuillage supplémentaire des niveaux de gris dans l'espace image permet de s'affranchir des derniers contours parasites (figure 4.8b).



Figure 4.8 : a) contour reconstruit après le seuillage dans l'espace de Hough, b) contour reconstruit après seuillage de l'image a.

Nous avons vérifié la validité de cette première étape de l'algorithme en comparant les valeurs des paramètres de droites obtenues par cet algorithme avec ceux calculés à la main. Les résultats sont donnés dans le tableau 4.1 pour le contour de la figure 4.6a.

a = -1.69 b = 45.85	a = -1.71 b = 46.60
a)	b)

Tableau 4.1 : vérification des valeurs des coefficients de la droite pour le contour considéré – a) valeurs calculées par l'algorithme, b) valeurs obtenues à la main.

Cinq autres contours ont été testés et le tableau 4.2 reporte les résultats pour les coefficients a et b calculés à la main et par l'algorithme. Nous avons aussi indiqué les erreurs moyennes absolues pour chacun des paramètres.

<i>Indice du contour</i>	<i>Paramètres calculés à la main</i>	<i>Calculés grâce à l'algorithme</i>	<i>Erreur moyenne absolue (MAE)</i>
<i>1</i>	a = -1.83 b = 40.45	a = -1.78 b = 40.02	MAE(a) = 0.05 MAE(b) = 0.73
<i>2</i>	a = -1.73 b = 25.81	a = -1.78 b = 26.12	
<i>3</i>	a = -1.83 b = 44.28	a = -1.80 b = 43.20	
<i>4</i>	a = -1.79 b = 28.50	a = -1.70 b = 27.44	
<i>5</i>	a = -1.71 b = 46.60	a = -1.69 b = 45.85	

Tableau 4.2 : résultats obtenus pour cinq autres contours.

On remarque que l'erreur (MAE) commise sur a est moins élevée que celle sur le paramètre b qui reste toutefois inférieure au pixel. Pour assurer une bonne superposabilité des différentes versions du profil, le paramètre le plus important est la valeur de la pente de la droite. Grâce à ces résultats, nous pouvons conclure que nous avons implanté correctement le calcul de l'équation de la droite.

4.2.4 Du profil suréchantillonné à la FTM

La droite obtenue permet de localiser avec précision le contour. Ces coordonnées permettent de connaître en chaque ligne la position précise du point d'inflexion du profil et, ainsi, de pouvoir superposer ultérieurement les différentes versions du profil final. On définit alors un voisinage autour du point d'inflexion qui permettra d'extraire seulement la zone d'intérêt, c'est-à-dire une version du profil pour chaque ligne. Ce tube est défini par une largeur de plus ou moins n pixels autour du point d'inflexion. Alors, chaque lecture de ligne donne une version du profil comportant $2n+1$ éléments.

On fixe alors le facteur de suréchantillonnage du contour. Dans nos applications numériques, nous l'avons choisi égal à 2, ce qui nous permet de synthétiser une courbe *FTM* entre $[-fe, +fe]$. On définit la longueur du profil final par l'expression $2(2n+1)+1$ cases. On joint à ce profil un tableau qui sert de compteur au précédent. La case centrale de ce profil final correspond à la localisation du point d'inflexion. La première ligne extraite de l'image vient remplir les cases de ce profil en fonction de la distance de chacun des pixels qui la composent par rapport au point d'inflexion. On incrémente le compteur des cases correspondantes. Le second extrait de ligne peut alors soit tomber dans les mêmes cases que la première ligne et donc incrémenter les compteurs des mêmes cases que pour la première ligne, soit remplir les cases d'à côté puisque le point d'inflexion se déplace légèrement à cause de la pente de la transition. Les compteurs de ces cases, vides au départ, sont placés à 1. On réitère l'opération pour toutes les lignes de l'image. Les valeurs de chaque case du profil final sont divisées par celle du compteur correspondant, permettant ainsi de moyenner les valeurs du profil. De plus, celui-ci contient deux fois plus de valeurs que les extraits de ligne initiaux ; on obtient donc bien un profil suréchantillonné d'un facteur 2.

Pour éradiquer les dernières traces de bruit, on minimise la distance entre la version finale du profil obtenu avec une fonction analytique. La courbe la plus proche du contour recherché est la tangente hyperbolique. Ce type de profil s'apparente à celui de la réponse des neurones à un stimulus extérieur. La fonction est alors connue sous le nom de fonction sigmoïde ou de fonction logistique. Elle s'écrit :

$$f(x) = \frac{1}{a_0 a_1^x + a_2} \quad \text{Eq 4.17}$$

La figure 4.9a reprend le contour qui nous sert d'illustration depuis le début de la description de notre méthode. La courbe en trait plein située en b correspond à la version moyennée du profil final, lequel est donc suréchantillonné d'un facteur 2. Celle en pointillés donne le résultat de la minimisation avec la sigmoïde grâce à la fonction COMFIT d'IDL. Cette fonction IDL permet de minimiser par la méthode des moindres carrés la distance entre une série de points et quelques fonctions usuelles. Il suffit de préciser le nom de la fonction que l'on souhaite utiliser en spécifiant son nom en paramètre de la fonction.

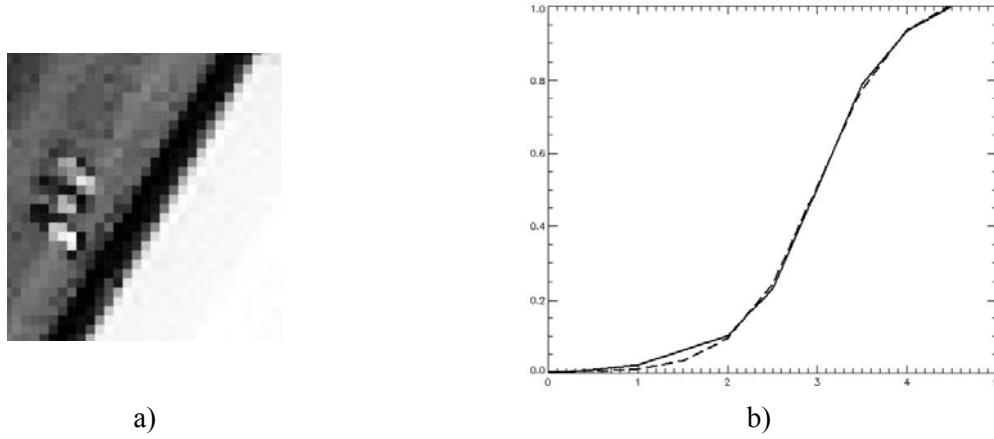


Figure 4.9 : a) rappel du contour traité, b) en trait plein, le contour suréchantillonné et moyenné. En pointillés, la sigmoïde correspondant le mieux au profil.

La norme $L2$ est calculée pour vérifier la performance de l'ajustement. Plus la valeur de la norme sera faible, plus la qualité de la minimisation est bonne. La norme $L2$ entre deux séries de m valeurs $s1$ et $s2$ s'écrit :

$$L2(s1, s2) = 1/m * \sqrt{\sum (s1(i) - s2(i))^2} \quad \text{Eq. 4.18}$$

La distance $L2$ est normée par rapport au nombre d'échantillons.

Il y a plusieurs avantages à manipuler une fonction analytique :

- ♦ sa dérivée est facilement calculable, et est aussi une fonction analytique qui peut être directement implantée avec les paramètres issus de la minimisation. Cette dérivée s'écrit :

$$f'(x) = \frac{-a_0 \ln(a_1) a_1^x}{(a_0 a_1^x + a_2)^2} \quad \text{Eq 4.19}$$

- ♦ le nombre d'échantillons peut parfaitement être contrôlé. Il s'agit d'un avantage considérable puisque la transformée de Fourier qui succède à la dérivation nécessite un nombre d'échantillons qui soit une puissance de 2,
- ♦ le point d'inflexion peut être théoriquement calculé par l'annulation de la dérivée seconde de la sigmoïde. Cette dérivée seconde s'écrit :

$$f''(x) = \frac{(a_0^2 a_1^{2x} - a_2^2)(a_0 \ln^2(a_1) a_1^x)}{(a_0 a_1^x + a_2)^4} \quad \text{Eq 4.20}$$

Elle s'annule pour x valant :

$$x_{\text{inflexion}} = \frac{\ln\left(\frac{a_2}{a_0}\right)}{\ln(a_1)} \quad \text{Eq 4.21}$$

La courbe de la figure 4.10 est le résultat obtenu pour le contour de la figure 4.9a. La valeur de la FTM à f_N vaut 0.11 et à la moitié de f_N vaut 0.48. Le suréchantillonnage par un coefficient 2 permet d'observer les fréquences au-delà des limites imposées par Shannon et la FTM se rapproche de la valeur 0 à f_e . Le repliement spectral est en partie éliminé par notre approche.

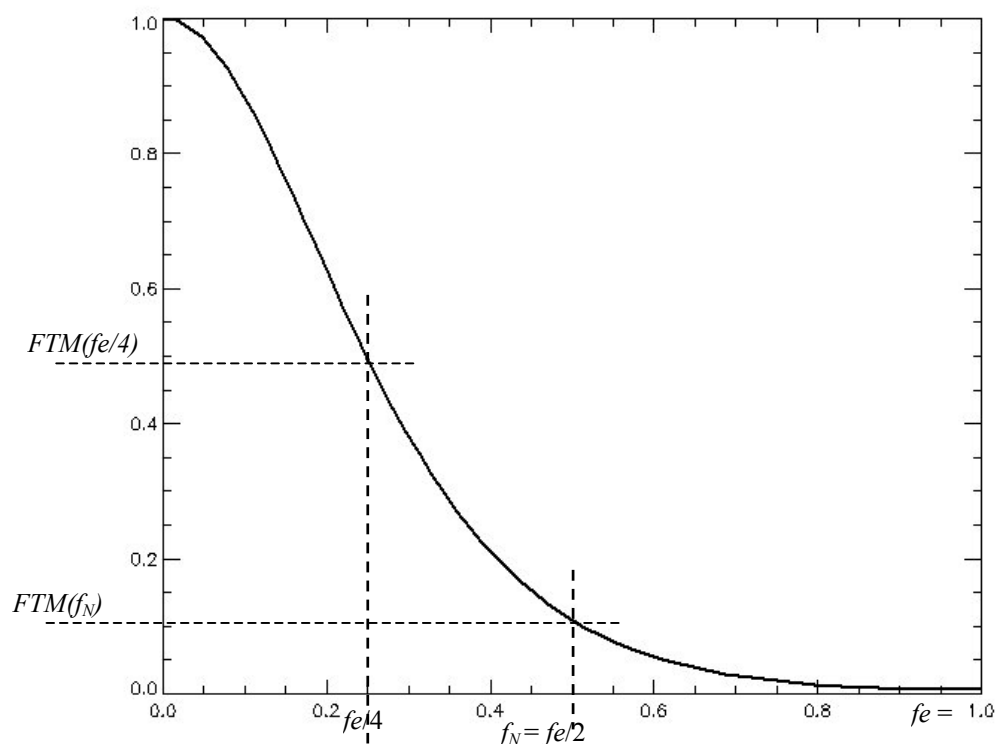


Figure 4.10 : résultat de la courbe obtenue pour le contour pour l'illustration.

4.2.5 Validation de l'algorithme

Pour la validation, nous nous sommes appuyés sur les travaux de Dial *et al.* (2003) et Ryan *et al.* (2003) pour respectivement les modalités Pan et bleue du satellite Ikonos. Ces auteurs ont déterminé la *FTM* des capteurs Ikonos à partir de cibles placées dans des champs dans le but de les comparer aux spécifications initiales. Nous avons procédé de la manière suivante : dix contours ont été sélectionnés dont cinq dans des images Pan et cinq dans des images bleues. Chaque contour donne une valeur de la *FTM* à la fréquence de Nyquist et nous comparons la moyenne de ces valeurs et leur dispersion pour ces deux modalités.

4.2.5.1 Ikonos Pan

La spécification pour la modalité Pan est de 0.17, ce qui correspond à la valeur de la *FTM* à la fréquence de Nyquist. Six estimations de la *FTM* sont disponibles dans l'article de Dial *et al.* (2003) ce qui permet d'avoir un intervalle autour de cette valeur clef : les valeurs données par les différentes estimations vont de 0.12 à 0.20.

Nous avons sélectionné cinq contours dans diverses images Pan d'Ikonos. La résolution spatiale de 1 m permet de sélectionner des bâtiments. La qualité des transitions est bonne en ce qui concerne la différence en amplitude et le bruit est limité. Par contre, la longueur des transitions est assez faible puisque peu de bâtiments ont une taille qui excède une trentaine de mètres.

Les résultats de nos différentes estimations sont présentés dans le tableau 4.3 comme suit : les cinq contours traités sont indicés de 1 à 5. Pour chacun, nous indiquons la longueur du contour en nombre de pixels, la taille de la transition en terme de différence de niveaux de gris, la norme *L2* qui indique la performance de la minimisation entre le profil suréchantillonné et moyenné et la sigmoïde et enfin la valeur de la *FTM* à la fréquence de Nyquist. Enfin, on reporte la valeur moyenne de *FTM* et les valeurs extrêmes.

<i>Indice du contour</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
<i>Longueur du contour (en nb de pixels)</i>	26	22	29	21	36
<i>Différence de radiométrie</i>	1500	1700	1550	1500	1500
<i>Norme L2</i>	0.025	0.007	0.001	0.006	0.009
<i>Valeur de la FTM à la fréquence de Nyquist</i>	0.318	0.309	0.138	0.190	0.133
<i>Valeur moyenne et intervalle</i>	0.22 [0.13, 0.32]				
<i>Valeur moyenne et intervalle sans le 1^{er} contour</i>	0.19 [0.13, 0.31]				

Tableau 4.3 : résultats de la routine d'estimation de la *FTM* pour la modalité Pan du satellite Ikonos

La valeur de la *FTM* moyenne estimée sur les cinq contours est légèrement surestimée par rapport à la valeur des spécifications (0.17). Or, on remarque que la qualité de la minimisation entre le profil du contour 1 et la sigmoïde est plus mauvaise d'un facteur 10, en ce qui concerne la norme *L2*, que pour les quatre autres contours. Si on disqualifie ce contour qui était excessivement bruité, la moyenne est ramenée à 0.19 et l'intervalle à [0.13, 0.31], ce qui cadre bien avec les valeurs attendues : 0.17 et [0.12, 0.20].

4.2.5.2 Ikonos modalité bleue

La cible artificielle utilisée dans l'article de Ryan *et al.* (2003) était un objet rectangulaire de 12 m sur 60 m, orienté de 8 degrés par rapport au nord, produisant un suréchantillonnage de 20. La spécification pour la *FTM* de la modalité bleue est de 0.26 à la fréquence de Nyquist et leurs six différentes estimations ont couvert des valeurs allant de 0.22 à 0.32. La sélection des contours a été limitée par la résolution spatiale des images bleue d'Ikonos. A 4 m de résolution spatiale, les bâtiments ne sont plus assez longs pour l'initialisation de l'algorithme. Les cinq contours extraits correspondent à des bords de routes ou à des ponts, puisque la résolution spatiale n'est plus que de 4 m. Contrairement à ceux extraits pour l'imagerie Pan, les contours vont pouvoir être beaucoup plus longs dès lors que l'on travaille sur des scènes aux Etats-Unis dont les routes ont la particularité d'être longues et rectilignes. Par contre, le bruit des zones qui jouxtent le contour est très important puisqu'il s'agit en général de parcelles agricoles. Le gradient de radiométrie est en général plus faible.

Le tableau 4.4 montre nos résultats en suivant la même organisation que le tableau précédent.

<i>Indice du contour</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
<i>Longueur du contour (en nb de pixels)</i>	212	170	67	162	218
<i>Différence de radiométrie</i>	800	700	1300	100	200
<i>Norme L2</i>	0.001	0.012	0.003	0.001	0.003
<i>Valeur de la FTM à la fréquence de Nyquist</i>	0.358	0.586	0.322	0.138	0.297
<i>Valeur moyenne et intervalle sans le 2nd contour</i>	0.28 [0.14, 0.36]				

Tableau 4.4 : résultats pour la modalité bleue Ikonos.

Le second contour a été directement exclu des conclusions. Le résultat de la minimisation montre que l'on n'a pas pu convenablement ajuster la fonction analytique. La *FTM* issue de ce contour est très largement au-dessus de la courbe de référence, ce qui explique la valeur importante de la *FTM* à la fréquence de Nyquist. La moyenne et l'intervalle sont proches des spécifications (0.26 et [0.22, 0.32]).

On pourrait envisager de procéder à nouveau à la validation sur les autres modalités Ikonos puisque Dial *et al.* (2003) donnent les valeurs des spécifications pour les autres modalités MS de ce

satellite (tableau 4.5). Mais, par manque de temps, nous n'avons pas eu l'occasion de le faire dans le cadre de cette thèse.

<i>Modalité</i>	<i>Pourcentage d'atténuation de la FTM à f_N</i>
Pan	0.17
bleue	0.26
verte	0.28
rouge	0.29
PIR	0.28

Tableau 4.5 : $FTM(f_N)$ pour les différentes modalités du satellite Ikonos (Dial *et al.* 2003).

4.2.6 Conclusion sur la méthode d'estimation et perspectives d'amélioration

Le but était d'éviter le recours à l'utilisation de cibles artificielles. Bien que l'utilisateur doive intervenir pour sélectionner un contour dans une image, il peut se satisfaire d'images qu'il a déjà à sa disposition. Cette méthode permet, à la condition qu'un contour d'intérêt soit identifié, de connaître la valeur de la FTM à la fréquence de Nyquist de cette image en particulier, puisqu'il est possible que deux images d'un même satellite donnent des FTM différentes. En effet, nous avons eu l'occasion de manipuler deux images Pan du satellite Ikonos qui ne donnaient pas la même FTM ; l'une des deux avait subi un rehaussement des contours par un filtrage passe-haut. Un autre avantage de la méthode présentée est sa simplicité d'implantation.

Notre méthode d'estimation de la courbe FTM se base sur l'hypothèse que le profil peut être ajusté par une fonction analytique qui est la sigmoïde. Cette supposition est importante car elle dicte la forme finale de la FTM . La FTM finale doit être représentative de la chaîne d'acquisition comme nous l'avons expliquée en première partie. Elle doit donc présenter une forme d'exponentielle décroissante (tâche d'Airy) multipliée par des sinus cardinaux (tâche d'Airy, électronique, effet de filé). L'exponentielle décroissante impose une forte pente décroissante à l'origine, or notre courbe présente une tangente nulle à l'origine. L'origine de ce défaut réside dans le choix de la fonction analytique. La TF de la dérivée de la sigmoïde ne permet pas d'obtenir une singularité en 0. On retrouve très souvent cette anomalie par exemple dans les méthodes qui prennent comme modèle de PSF une gaussienne. On remarque aussi l'absence dans notre modèle des rebonds provoqués par la présence des sinus cardinaux dans le modèle idéal.

On peut aussi reprocher l'initialisation manuelle de notre méthode. En effet, cette étape dépend de l'utilisateur, elle est donc totalement subjective. Nous proposons une autre approche pour pallier cette limitation ; l'utilisation d'une transformation en curvelettes. En effet, au contraire des bases d'ondelettes qui ne permettent pas de détecter des formes allongées et/ou orientées, les curvelettes sont capables de détecter des objets présentant une régularité plus forte dans une direction donnée. Dans notre cas, le contour présentera une régularité importante selon sa tangente et très faible selon la direction normale au contour.

4.3 Proposition d'une nouvelle distance pour l'évaluation de la qualité géométrique des produits de fusion (L_2CFC)

La méthode présentée est un bon estimateur de la valeur de la FTM à la fréquence de Nyquist d'un capteur si suffisamment de contours ont été sélectionnés dans une ou plusieurs images d'un même satellite et si on a l'assurance qu'aucun post-traitement n'a été appliqué à l'une des images. Mais la méthode peut être exploitée de manière différente ; à partir d'une transition d'une image acquise par un capteur, on en déduit une information en fréquence qui témoigne de la qualité du capteur. Toutefois, si le contour d'une image est désignée pour être la référence, on connaîtra les fréquences spatiales qui le composent ainsi que leur proportion, en fonction du bruit et du gradient de sa transition. Voilà pourquoi nous avons eu l'idée d'utiliser cet outil pour caractériser la qualité des

produits de fusion. Il est possible que les méthodes de fusion distordent localement ou de manière globale les contours d'une image fusionnée. Ainsi, si on applique notre approche à un contour de l'image de référence, on obtiendra les informations en fréquences de ce contour, qui doivent être identiques à celles de l'image fusionnée. Plus la distance entre les courbes fréquentielles obtenues sera grande, plus l'algorithme de fusion aura rencontré des difficultés à synthétiser ce contour. On ne peut plus à proprement parler de *PSF* ni de *FTM* puisqu'une image fusionnée n'est pas acquise par un capteur. Ces termes permettent de caractériser les faiblesses et performances d'un capteur. On parlera plutôt de qualité géométrique locale ou de contenu fréquentiel du contour (*CFC*). Soit L_2CFC cette distance.

Cet outil a été testé sur deux extraits issus des simulations CNES, déjà décrites et utilisées dans le chapitre 3 consacré à la vérification de l'hypothèse d'extrapolation. Les références disponibles aux diverses résolutions spatiales nous ont permis d'évaluer la qualité géométrique des images résultantes de plusieurs méthodes de fusion.

4.3.1 Extrait d'une arête d'un immeuble de Toulouse

Ce premier exemple nous permet de décrire la manière d'utiliser l'outil. Nous avons tout d'abord appliqué trois méthodes de fusion sur un premier extrait dont nous connaissons *a priori* les artefacts géométriques produits sur les contours des images fusionnées. Ces algorithmes appartiennent au concept *ARSIS*. Ce concept applique des représentations multiéchelles aux images d'entrée dans le but d'estimer un modèle pour transformer les détails de l'image Pan vers l'image MS. Certaines de ces représentations sont connues pour distordre la qualité géométrique des produits de fusion.

L'extrait de référence se trouve figure 4.11a. La première méthode de fusion utilise l'algorithme de décomposition multiéchelle 'à trous' *ATWT* (« *A trous* » *Wavelet Transform*) associé à une estimation locale des paramètres, appelée *RWM* que nous décrirons dans le chapitre suivant (figure 4.11b). La qualité géométrique atteinte par cet algorithme peut soit être localement médiocre, soit être localement de très bonne qualité. Nous avons sélectionné un extrait où le contour a correctement été traité par l'algorithme de fusion. La seconde méthode de fusion utilise le même algorithme de décomposition *ATWT* et utilise un modèle global d'estimation des paramètres appelé *M3* (Ranchin *et al.* 1997) décrit dans le chapitre suivant (figure 4.11c). La qualité atteinte par cette méthode de fusion est identique sur toute l'image mais l'image résultante est généralement jugée trop floue. Enfin, la dernière méthode de fusion repose sur une analyse multirésolution telle qu'elle a été développée par *Mallat*, qui utilise une transformée en ondelettes définie par Daubechies, associée au modèle global *M3*. La classe d'ondelettes définie par Daubechies comprend des membres couvrant une large palette en terme de localisation et de régularité. La plus simple et la plus localisée, est celle que nous avons utilisée, souvent appelée *DAUB4* car elle possède quatre coefficients. L'ondelette *DAUB4* n'est pas symétrique, et au contraire de la pyramide Laplacienne, introduit un déphasage d'un demi pixel par échelle de décomposition. Lorsque la modalité Pan est dégradée deux fois pour atteindre la résolution spatiale de la modalité à traiter, les images ne sont plus superposables. L'effet est dévastateur sur l'image fusionnée comme en témoigne la figure 4.11d.

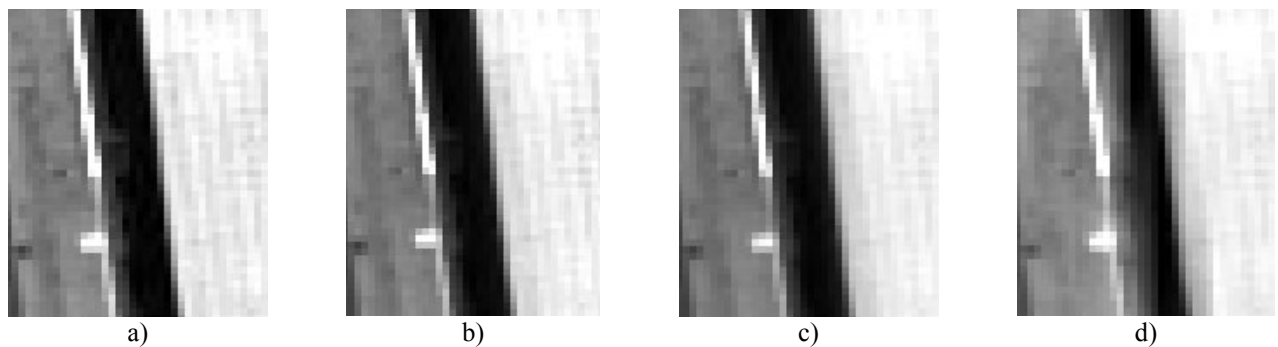


Figure 4.11 : simulation CNES sur la ville de Toulouse. a) contour de référence, b) contour synthétisé par *ATWT-RWM*, c) *ATWT-M3*, et d) contour obtenu par *MallatDaub4-M3*. Copyright CNES 2000.

Les courbes *CFC* pour chacun de ces trois contours sont représentées sur la figure 4.12a. La courbe noire correspond à celle de référence. Ensuite de la plus proche à la plus éloignée, nous avons les courbes bleue, verte et rouge représentant respectivement les résultats pour les contours b, c et d de la figure 4.11. Sur la figure 4.12b sont indiquées les valeurs obtenues par la distance L_2 entre chaque méthode de fusion et la référence. La norme L_2 est ramenée au nombre d'échantillons et représente donc l'écart moyen entre les points de la courbe de référence et ceux estimés, et tient donc compte des unités et de la dynamique. Ici, les valeurs L_2 sont très faibles, nous y apportons une précision de quatre chiffres après la virgule.

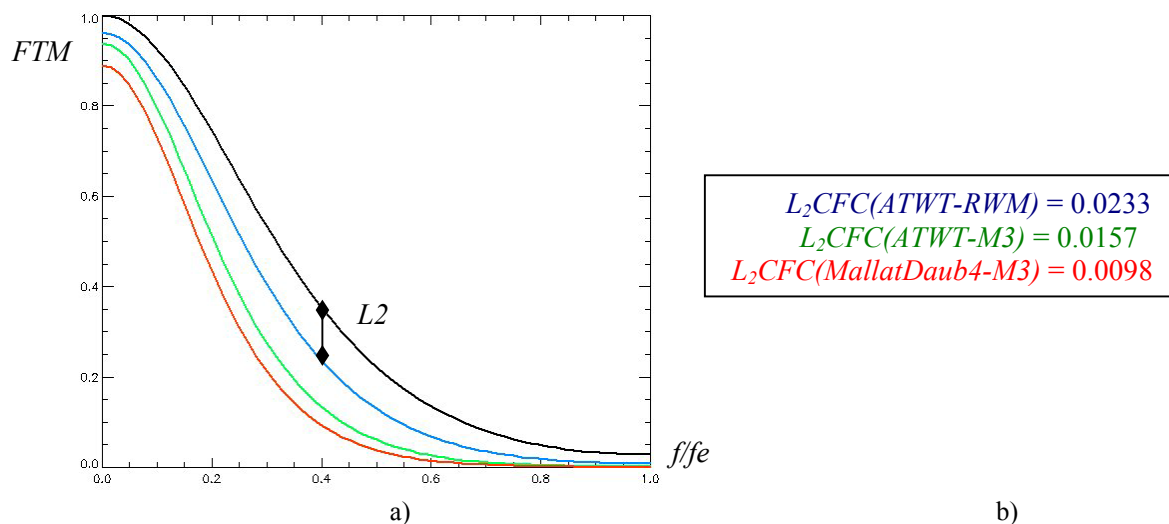


Figure 4.12 : a) courbes *CFC* pour la référence (courbe noire), pour le contour fusionné par *ATWT-RWM* (courbe bleue), par *ATWT-M3* (courbe verte) et enfin par *MallatDaub4-M3* (courbe rouge), b) à droite, valeurs de la norme L_2CFC entre la référence et chacune des méthodes de fusion.

La référence majore les résultats pour cet extrait. Les valeurs *CFC* des courbes fusionnées sont divisées par la valeur à l'origine de la référence. Cette sous-estimation signifie que les méthodes de fusion n'ont pas introduit de fréquences spatiales en quantité suffisante pour reproduire de manière fidèle le contenu fréquentiel du contour de référence.

L'interprétation des courbes et des valeurs des normes L_2CFC confirme les connaissances a priori sur la qualité des contours synthétisés par ces différentes méthodes. En effet, la courbe bleue de la méthode *ATWT-RWM* est bien la plus proche de la référence et la valeur de la norme L_2CFC entre la courbe bleue et la courbe noire est bien la plus faible des trois valeurs obtenues. La méthode de fusion basée sur l'utilisation d'une ondelette non-symétrique minore la qualité obtenue en terme de contenu fréquentiel du contour puisque sa courbe est la plus éloignée de la référence. Le contour a été détérioré ; il apparaît plus progressif que pour les autres méthodes de fusion. En terme de fréquence, cela se manifeste par une atténuation plus prononcée des fréquences spatiales. L'écart donné par les distances L_2CFC est en accord avec cette observation en donnant une valeur plus élevée que pour les autres.

En conclusion, les résultats apportés par cet outil semblent convaincants. Nous avons dès lors testé cinq méthodes de fusion relevant du concept *ARSIS*, dont les trois de l'exemple précédent. Nous avons souhaité voir l'influence du ratio sur les résultats des courbes *CFC*. Nous présentons tout d'abord les courbes de résultats pour le ratio 2 et ensuite pour le ratio 4. Une dernière partie est consacrée à l'étude de l'influence de l'algorithme ayant servi à la dégradation des MS à l'origine de la fusion. En effet, pour le ratio 4, au lieu de prendre les MS simulées par le CNES à 160 cm, nous avons utilisé un autre algorithme. Nous comparons ces résultats avec ceux obtenus avec l'algorithme de simulation du CNES pour le ratio 4.

4.3.2 Extrait de la prison de Toulouse

L'extrait sélectionné pour ces trois expériences correspond à un bord d'un des bâtiments de la prison de Toulouse (figure 4.13a). L'image d'origine de cet extrait est la modalité PIR de la simulation CNES à 40 cm. La courbe de la figure 4.13b correspond au profil de ce contour, moyenné sur la longueur du contour et suréchantillonné d'un facteur 2.

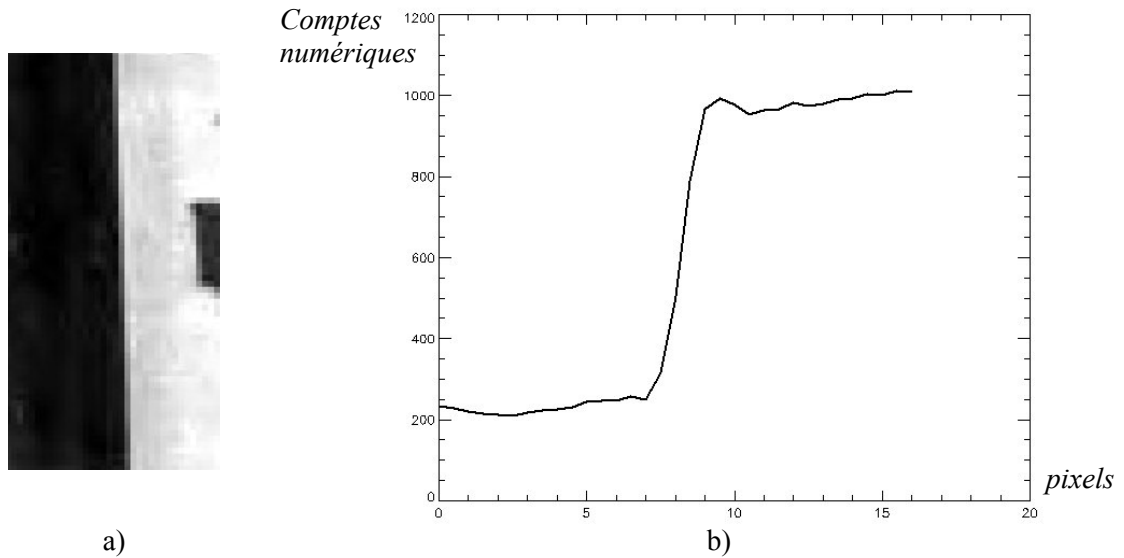


Figure 4.13 : extrait d'un des bâtiments de la prison de Toulouse. Copyright CNES 2000.

L'ajustement à une sigmoïde a donné une distance $L2$ normalisée de 0.0044. La figure 4.14 montre cette sigmoïde ainsi que le contour de référence.

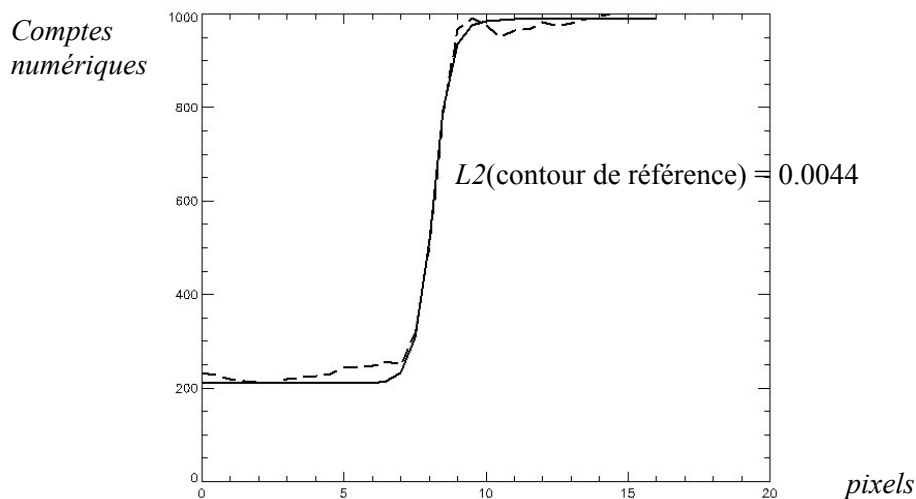


Figure 4.14 : contour de référence (en trait plein) et la fonction sigmoïde ajustée.

Les cinq méthodes de fusion testées sont les suivantes :

- ◆ *MallatDaub4-M3*,
- ◆ *ATWT-M3*,
- ◆ *ATWT-RWM*,
- ◆ *GLP-M3*,
- ◆ et *GLP-CBD*.

Leurs courbes *CFC* sont respectivement de couleur verte, rouge, bleue, vert pâle et violet foncé.

Deux méthodes de fusion ont été ajoutées par rapport à l'exemple précédent. Ces méthodes reposent sur le modèle de décomposition multiéchelle *GLP* correspondant à la pyramide Laplacienne

généralisée, avec respectivement le modèle global d'estimation $M3$ pour la première des deux méthodes, et un modèle d'estimation local (comme RWM) basé sur les contextes CBD pour *Context-Based-Decision* pour la deuxième méthode. Une campagne d'interprétation de la qualité visuelle de ces cinq produits de fusion a été organisée pour affiner notre connaissance des artefacts engendrés par ces méthodes.

4.3.2.1 Ratio 2

Pour synthétiser des images avec un ratio 2, nous avons choisi des images MS à 40 cm pour les références, et donc celles à 80 cm pour servir d'images originales pour la fusion. Les images fusionnées ont alors une résolution de 40 cm.

Les courbes CFC pour ce contour particulier et pour un ratio de fusion de 2 sont reportées figure 4.15 avec les valeurs des normes L_2CFC .

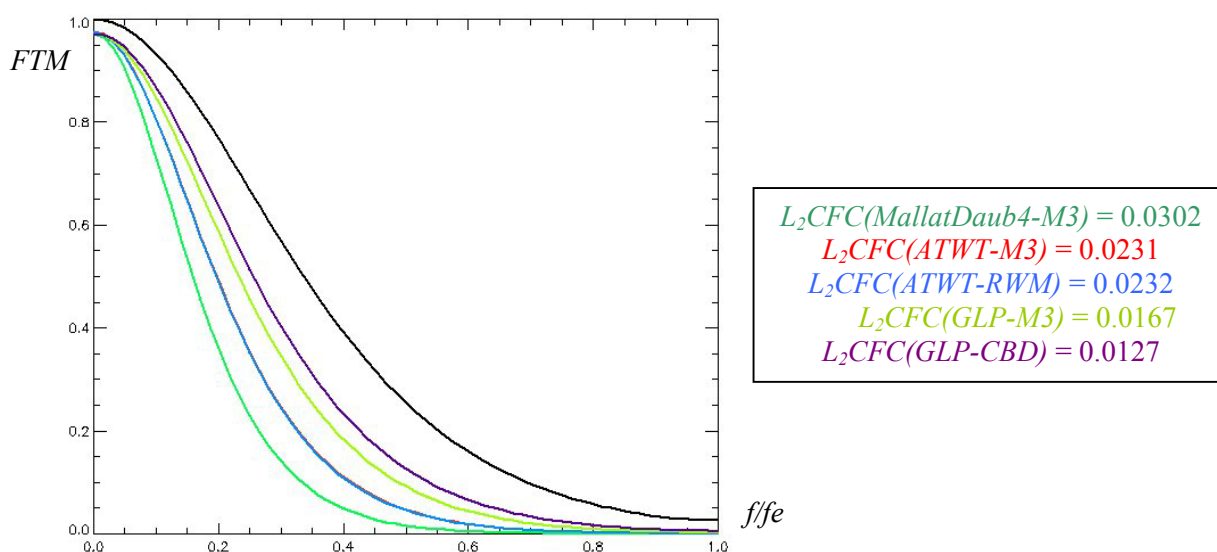


Figure 4.15 : ratio 2 - courbes CFC pour la référence (courbe noire) et les cinq contours fusionnés et valeurs de la norme L_2CFC entre la référence et chacune des méthodes de fusion.

Une fois encore, la référence est un majorant des résultats pour cet extrait. La méthode de fusion la moins performante se trouve être une fois encore la méthode utilisant l'algorithme de *MallatDaub4*. Les courbes rouge et bleue, correspondant respectivement aux méthodes de fusion *ATWT-M3* et *ATWT-RWM* sont superposées. Pour cet exemple, la qualité du contour obtenue par l'algorithme à trous est pratiquement identique, que l'on considère le modèle global $M3$ ou le modèle local RWM . La meilleure qualité géométrique rapportée par cet outil semble être la méthode de fusion *GLP-CBD*. En effet, sa courbe est la moins éloignée de la référence. Pour cet algorithme de décomposition, la qualité obtenue avec le modèle global $M3$ est légèrement moins bonne que celle obtenue avec CBD . Ces observations sont en adéquation avec celles faites pour les méthodes *ATWT-M3* et *ATWT-RWM*, à savoir qu'un modèle d'estimation local semble être plus performant pour reproduire les transitions des images. En conclusion, les méthodes de fusion utilisant le *MSM GLP* semblent être les plus performantes.

Lors de ces expériences, nous avons noté l'importance de la qualité de l'ajustement dans le résultat. Nous avons sélectionné un contour de référence qui s'adapte visuellement bien à un profil sigmoïdal. La distance $L2$ obtenue entre les points du contour de référence et ceux de la sigmoïde qui a été ajustée (figure 4.14) est de 0.0044. Les valeurs $L2$ placées figure 4.16 représentent les résultats des ajustements des sigmoïdes pour les différentes méthodes de fusion. On remarque que ces valeurs sont toutes supérieures à celles obtenues pour l'ajustement du contour de référence. Tout comme la figure 4.14, nous avons placé à droite de la figure 4.16 le contour ainsi que la sigmoïde ajustée pour la méthode de fusion *GLP-CBD*, car il s'agit de celle dont la distance $L2$ était la plus élevée. Comparée à

l'ajustement sur le profil de référence, on note que les points du profil obtenu pour cette méthode de fusion s'éloignent de la sigmoïde dans les deux régions de part et d'autre de la pente.

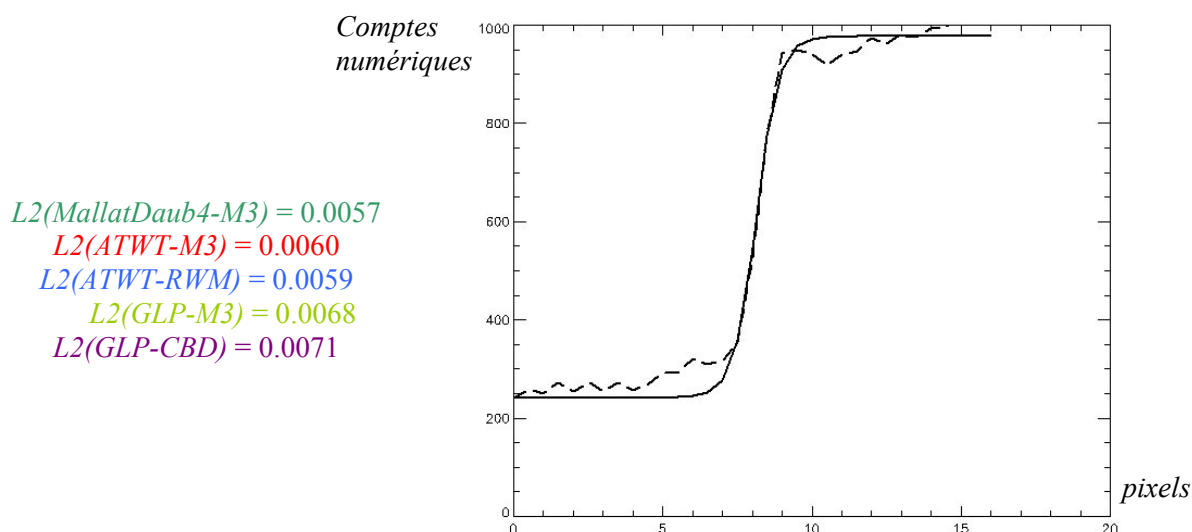


Figure 4.16 : résultats de l'ajustement par une fonction sigmoïde pour les cinq méthodes de fusion différentes et minimisation obtenue pour la distance L_2 la plus importante obtenue pour *GLP-CBD*.

Ainsi, l'ajustement a rencontré plus de difficultés pour les contours obtenus par les diverses méthodes de fusion que pour le contour de référence. Il semble judicieux de s'interroger sur l'influence de la qualité de l'ajustement sur la fiabilité des courbes *CFC* obtenues.

4.3.2.2 Ratio 4

Pour synthétiser des images avec un ratio 4, nous avons choisi des images MS à 160 cm simulées par le CNES pour servir d'images originales pour la fusion, et les images fusionnées possèdent alors la résolution du Pan qui est de 40 cm, où la référence MS est disponible.

Les courbes *CFC* pour la référence et les cinq contours fusionnés avec un ratio de fusion de 4 sont reportés sur la figure 4.17. Comme précédemment, les normes L_2 entre les courbes *CFC* de référence et fusionnées sont visibles à la droite du graphique.

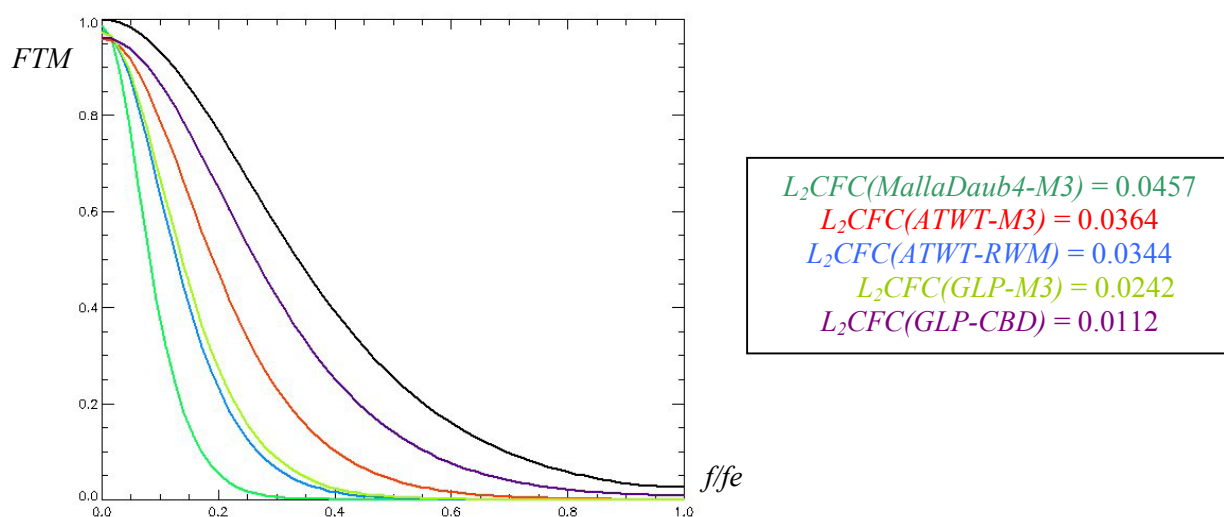


Figure 4.17 : ratio 4 - courbes *CFC* pour la référence (courbe noire) et les cinq contours fusionnés et, à droite, valeurs de la norme L_2 entre la référence et chacune des méthodes de fusion.

La référence reste un majorant de toutes les courbes *CFC* obtenues. Les résultats sont moins bons pour le ratio 4 que pour le ratio 2. L'ordre des courbes est relativement le même que pour le ratio

2, avec toutefois une inversion des tendances entre le résultat pour *ATWT-M3* et *ATWT-RWM*. Cependant, les conclusions doivent être relativisées car les difficultés d'ajustement ne sont, encore une fois, pas prises en compte dans ce résultat. En effet, les valeurs $L2$ obtenues pour quantifier la réussite de l'ajustement (figure 4.18) sont très élevées, excepté pour la première méthode de fusion *MallatDaub4-M3*.

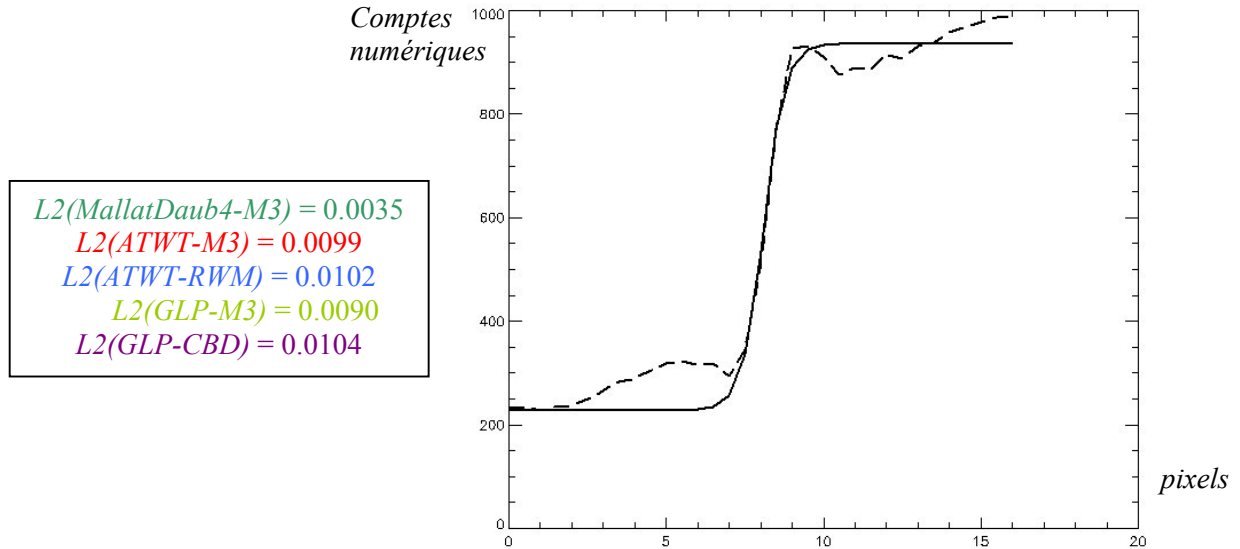


Figure 4.18 : ratio 4 - résultats de l'ajustement par une fonction sigmoïde pour les cinq méthodes de fusion différentes, et ajustement obtenu pour la distance $L2$ la plus importante obtenue pour *GLP-CBD*.

On peut s'attendre à ce que la courbe *CFC* pour la méthode *Mallat-M3* soit une bonne estimation du résultat puisque la distance $L2$ pour minimisation est même plus faible que pour la référence. Par contre, l'erreur $L2$ d'ajustement pour les autres méthodes de fusion est environ deux fois plus importante que celle commise sur le contour de référence. La courbe placée en b correspond à la distance $L2$ de *GLP-CBD*, car elle présente encore une fois la distance $L2$ la plus élevée. On remarque ici que la forme de la sigmoïde n'est pas forcément la mieux adaptée au profil obtenu par *GLP-CBD*, puisque excepté le long de la pente, les points du profil ne présentent pas une pente nulle, contrairement à ceux de la sigmoïde. La qualité des estimations des courbes *CFC* des autres méthodes de fusion doit être relativisée car ces estimations sont entachées des erreurs d'ajustement.

4.3.2.3 Ratio 4 – synthèse des MS à 160 cm par GLP

Certains auteurs déplacent le problème dans les échelles parce que par exemple, ils disposent d'images MS seulement à une résolution (Munehika *et al.* 1993, Wald *et al.* 1997, Thomas et Wald 2005). Cette manière de procéder permet de s'affranchir du problème de la référence. Nous avons aussi recommandé ce changement d'échelle dans la vérification des deux propriétés du protocole. Nous avons voulu tester ici l'influence de l'algorithme de dégradation sur la qualité des résultats. Pour ce faire, nous avons synthétisé les MS à la basse résolution spatiale de 160 cm à partir de celles à 40 cm que nous avons volontairement dégradées grâce à l'algorithme de dégradation *GLP*. Notons que cet algorithme est le même que celui employé dans les méthodes de fusion *GLP-M3* et *GLP-CBD*. Les courbes *CFC* et valeurs de $L2$ sont placés figure 4.19.

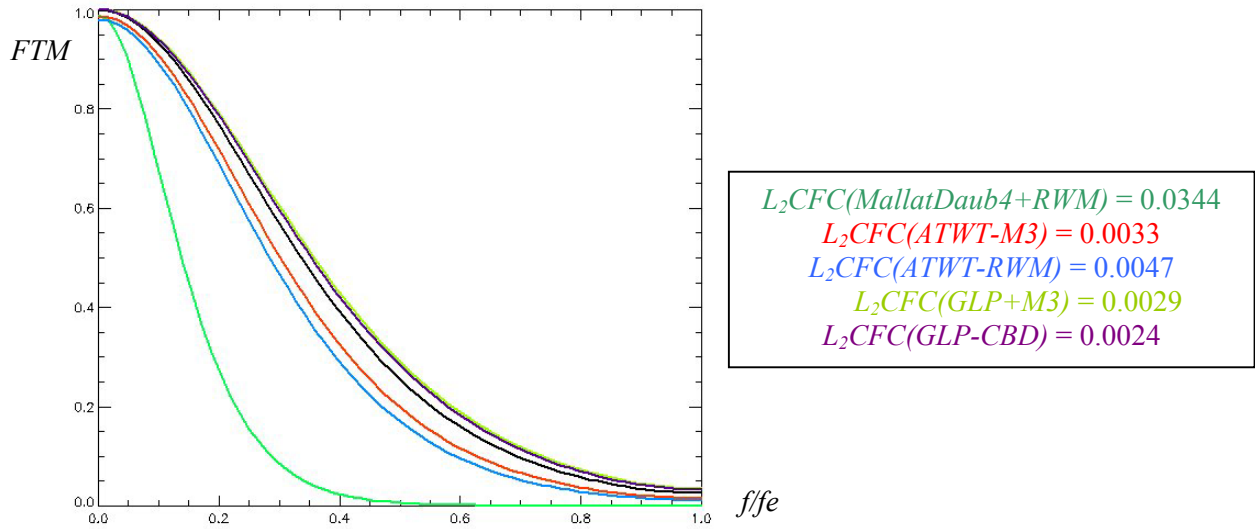


Figure 4.19 : ratio 4 – images MS à 160 cm obtenues par *GLP* - courbes *CFC* pour la référence (courbe noire) et les cinq contours fusionnés et valeurs de la norme L_2CFC entre la référence et chacune des méthodes de fusion.

A part celle de *MallatDaub4-M3* qui demeure très éloignée de la référence, les autres courbes sont nettement plus réunies autour de la courbe de la référence que dans le cas précédent où les images MS d'origine n'avaient pas été synthétisées par *GLP*. Pour ces quatre méthodes de fusion, la création des MS à 160 cm a été favorable en terme de qualité géométrique. L'algorithme de décomposition a donc une influence non négligeable sur le résultat. Deux groupes se distinguent : celui qui utilise l'algorithme *ATWT* et celui de l'algorithme *GLP*. Ce sont les méthodes utilisant *GLP* qui ont été le plus favorisées ; la qualité obtenue surévalue même celle du contour d'origine. De même, celles utilisant *ATWT* ont eu moins de difficultés dans la fusion. Les valeurs des distances L_2 du succès de l'ajustement sont plus faibles (figure 4.20).

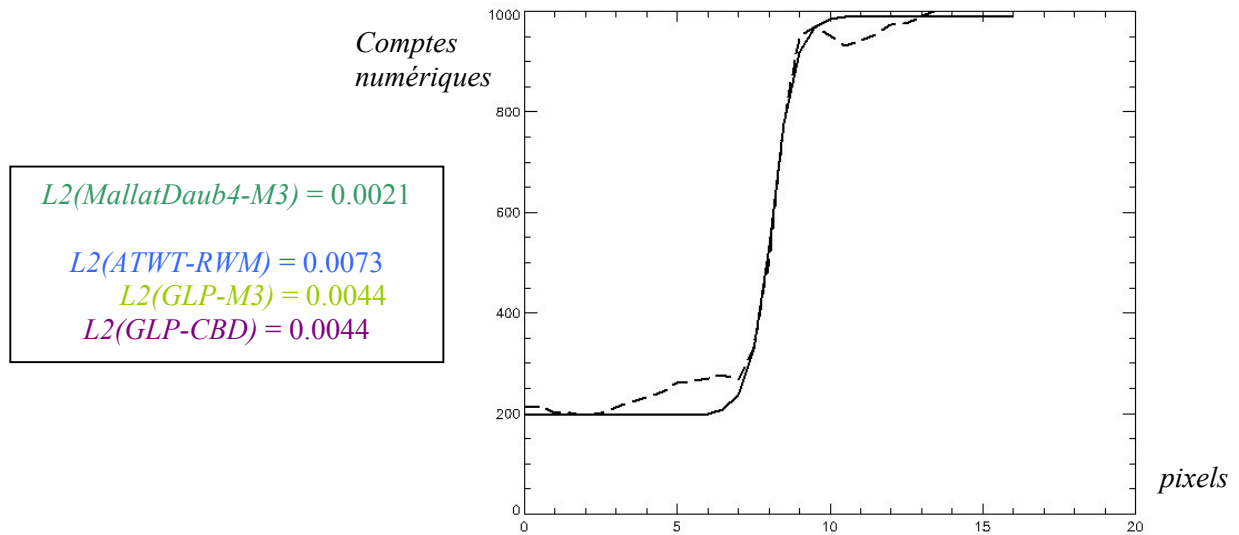


Figure 4.20 : ratio 4 – MS simulées à basse résolution par *GLP*. résultats de l'ajustement par une fonction sigmoïde pour les cinq méthodes de fusion différentes, et ajustement obtenu pour la distance L_2 la plus importante obtenue (*ATWT-RWM*).

Nous concluons que les performances d'une méthode de fusion dépendent fortement des conditions d'acquisition et/ou de synthèse des modalités d'entrée. Les résultats sont moins bons lorsque les modalités MS sont simulées avec la chaîne de simulation du CNES que dans le cas où elles sont synthétisées par *GLP*. Ces résultats sont en accord avec les résultats donnés par le BQ monomodal de la modalité PIR synthétisée par les différentes méthodes de fusion. On rappelle que les

distances monomodales sélectionnées sont : $diffVarRel$, σRel , cc , et $ccHF$. Les résultats statistiques de toutes les méthodes de fusion, y compris celles de *MallatDaub4-RWM*, sont parfaitement en accord avec les résultats obtenus, à savoir que la qualité est meilleure pour les images MS simulées à basse résolution avec l'algorithme *GLP*. Par exemple, si l'on choisit la méthode de fusion *ATWT-M3*, les valeurs respectives des distances obtenues pour ces différentes distances sont, pour les images simulées par le CNES de : 20 %, 10 %, 0.989 et 0.985. Ces valeurs deviennent pour les modalités simulées par *GLP* : 4 %, 5 %, 0.996 et 0.985.

D'après le chapitre précédent, nous savons que l'on peut avoir une idée de la difficulté de synthèse en appliquant des distances qui informent sur la difficulté de synthèse, comme par exemple $diffVarRel(A0, A2)$ et $diffHRel(A0, A2)$. Nous avons appliqué ces distances entre les modalités PIR de référence à 40 cm et celle simulée par le CNES à 160 cm, puis entre la PIR de référence et celle dégradée à 160 cm par le *GLP*. Il s'avère que la valeur de la différence de variance relative est moitié moindre dans le cas des modalités MS synthétisées à 160 cm par *GLP* (13 % contre 27 % dans le cas de la simulation PIR du CNES), ce qui signifie que la quantité d'information à introduire pour obtenir un produit de fusion à 40 cm dans le cas d'une simulation avec *GLP* est moins importante. Cette tendance est confirmée par les valeurs de différence d'entropie relative (0.89 % contre 1.56 % pour les simulations CNES).

L'adéquation entre la qualité géométrique rapportée par notre outil de qualité basé *FTM* et la qualité visuelle rapportée par la première campagne d'évaluation nous permet d'envisager d'inclure cette distance dans le bilan de qualité monomodal.

4.4 Conclusion

Nous avons développé un outil capable de donner la valeur de la *FTM* à la fréquence de Nyquist pour un capteur directement à partir de ses images. L'avantage essentiel est d'éviter d'avoir recours aux cibles artificielles, et de bénéficier d'une caractéristique importante qui décrit le comportement spectral des capteurs spatiaux.

Nous avons exploré une autre application de cet outil basé *FTM* afin d'y percevoir un moyen de quantifier la qualité géométrique d'une image par rapport à une image de référence. On vient extraire une information en fréquence d'un contour de l'image de référence. Si cette image subit un traitement quelconque, il se peut que le contour soit affecté et que son information en fréquence ait été altérée. Dans le cadre de la fusion d'images, cet outil permet de comparer le contenu fréquentiel d'un contour, c'est-à-dire sa courbe *CFC*, par rapport au contenu en fréquence de ce même contour obtenu par une méthode de fusion. Une norme a été définie afin de donner en une unique valeur une idée de la dégradation géométrique locale générée par les méthodes de fusion.

Grâce à des connaissances a priori sur la qualité géométrique des produits de fusion, on a pu vérifier la validité de la démarche et des résultats rapportés par l'outil. Une fois ce stade dépassé, nous avons testé plusieurs stratégies de fusion de type *ARSIS*, et fait varier deux paramètres, qui sont le ratio entre images Pan et MS, et l'algorithme de simulation des MS à basse résolution spatiale. Ces tests ont révélé, en accord avec les mesures statistiques, que la qualité géométrique est meilleure pour un ratio plus faible (2 plutôt que 4) et que l'influence de la dégradation ne devait pas être négligée. Lorsque la quantité de détails à introduire dans l'image est plus importante comme dans le cas des simulations CNES par rapport à l'algorithme *GLP*, les algorithmes rencontrent plus de difficultés au moment de la fusion. Cette quantité d'information peut par exemple s'exprimer par la différence de variance relative entre les plans d'approximation à haute et basse résolution spatiale.

Une autre conclusion de ce travail concerne l'influence de l'algorithme de dégradation spatial recommandé par le protocole pour le changement d'échelle. Si l'algorithme utilisé pour la simulation est identique à celui utilisé dans la fusion, comme dans notre cas pour *GLP-M3* et *GLP-CBD*, la fusion sera favorisée et les résultats auront tendance à être surévalués. Ce constat rejoint celui fait récemment

par Alparone *et al.* (2006). De manière générale, pour être comparables, que les bilans de qualité soient visuels ou statistiques, mono- ou multimodaux, ils doivent être établis dans les mêmes conditions expérimentales pour toutes les méthodes de fusion. De plus, les algorithmes de pré-traitements doivent être choisis de telle sorte que l'impartialité sur les résultats soit assurée. Comme les résultats de cet outil *FTM* sont en accord avec les résultats des autres bilans de qualité, nous proposons d'ajouter cette distance monomodale dans le bilan de qualité que nous recommandons.

Cependant, nous avons noté la nécessité de prendre en compte la qualité de l'ajustement sur le résultat final. De plus amples recherches et tests devraient être entrepris afin de définir par exemple un critère d'exclusion des contours en cas d'échec lors de l'ajustement. Si ce contour a été trop détérioré par la fusion ou si le profil obtenu est trop éloigné du profil analytique d'une sigmoïde, la distance $L2$ résultante est élevée et on pourrait alors déclarer la méthode de fusion comme ayant échoué à synthétiser ce contour particulier. On pourrait aussi envisager une autre possibilité, comme de pondérer le résultat final par la qualité de l'ajustement en donnant par exemple une courbe *CFC* ainsi qu'un intervalle composé de deux courbes *CFC* entre lesquelles la probabilité d'obtenir une bonne estimation de la courbe *CFC* est suffisante, c'est-à-dire supérieure à un seuil donné.

V Catégories de méthodes de fusion

5.1 Introduction

Pohl et Van Genderen (1998) ont proposé une classification des méthodes basée sur le niveau de décision, que l'on pourrait apparenter à une classification qui distingue les méthodes dites de *bas niveau* des méthodes de plus *haut niveau* : pixel, structure et enfin, identification. Les deux catégories dites de plus haut niveau correspondent à l'extraction et au traitement d'objets extraits de l'image. La catégorie *identification* apporte un niveau de connaissance supplémentaire par rapport au niveau *structure* puisque les objets extraits ont été identifiés.

Wang *et al.* (2005) ont plutôt cherché à montrer que toutes les méthodes de fusion pouvaient rentrer dans un cadre de travail unique, qu'ils ont appelé *GIF* pour *General Image Fusion*. Notons que cette formulation n'est pas nouvelle puisqu'elle a déjà été proposée sous forme d'une méthode de fusion par Price en 1999. Ils expriment la fusion d'images par une expression additive en partant du principe que de l'information doit être ajoutée aux images MS originales. Si $(A_l)^{interp}_0(i, j)$ représente la modalité A_0 haute résolution dégradée et sous-échantillonnée à la résolution des modalités B_{kl} , puis suréchantillonnée pour que ses dimensions s'adaptent à celles de A_0 , alors la représentation GIF se résume à :

$$B_{k0}^* = (B_{kl})^{interp}_0 + \alpha_{A_l, B_{kl}} (A_0 - (A_l)^{interp}_0) \quad \text{Eq. 5.1}$$

où la $k^{\text{ième}}$ modalité fusionnée peut s'écrire comme la somme de deux termes. Les opérations s'effectuent pixel à pixel. Le premier terme correspond à la $k^{\text{ième}}$ modalité originale rééchantillonnée pour atteindre la taille de A_0 . On y ajoute un second terme qui représente une pondération $\alpha_{A_l, B_{kl}}$ de la différence d'information contenue entre la modalité Pan et sa version dégradée A_l rééchantillonnée. Cette pondération peut s'avérer être une simple constante, jusqu'à une expression complexe des ensembles originaux. Wang et ses collaborateurs déterminent la performance des méthodes en étudiant la manière dont est calculé A_l et le coefficient de pondération α . Ce formalisme me gêne car ici le point de départ de la fusion est le terme $(B_{kl})^{interp}_0(i, j)$ au lieu de B_{kl} . La manière de suréchantillonner une image n'est pas unique. De plus, elle n'est pas systématiquement la première étape commune à toutes les méthodes de fusion.

Une autre critique est le formalisme de l'injection des structures dans l'image $(B_{kl})^{interp}_0$. En effet, une simple différence $(A_0 - (A_l)^{interp}_0)$ ne peut pas représenter les différentes méthodes qui suivent une approche multirésolution ou multiéchelle présentes dans la littérature.

Afin de dresser un état de l'art des méthodes de fusion existantes, nous avons choisi d'adopter la classification de Wald (2000). Trois types de méthodes ont été identifiés :

- ♦ projection-substitution,
- ♦ contribution spectrale relative,
- ♦ et enfin les méthodes de type *ARSIS*.

La première catégorie a été originellement proposée par Shettigara en 1992 sous le nom de *COS* pour les techniques de *COmponent Substitution* et sera présentée dans la partie 5.2. La partie 5.3 décrit les méthodes de type contribution spectrale relative, en partant de leur écriture générale jusqu'aux implantations particulières de *Brovey*, *Color Normalized* et *P+XS*. La partie 5.4 introduit le concept *ARSIS* et quelques-unes de ses diverses implantations. On clôt ce chapitre (partie 5.5) par une démonstration qui montre que seules les méthodes de fusion qui suivent une approche multiéchelle sont capables de produire des images synthétiques qui, une fois ramenées à la résolution de l'image

MS originale, respectent le contenu fréquentiel de l'image originale, c'est-à-dire assure une bonne cohérence avec l'ensemble MS original (propriété de cohérence).

5.2 Les méthodes de type projection-substitution (COS)

Les méthodes de fusion de type projection-substitution suivent une approche vectorielle, puisque toutes les modalités MS sont synthétisées simultanément dans le processus de fusion. Une image d'une scène est constituée d'enregistrements suivant différents canaux spectraux, caractéristiques du satellite d'observation (Ikonos, SPOT, ...). Les images de la scène ont des ressemblances plus ou moins importantes selon qu'elles contiennent ou non les mêmes informations. L'objectif de la projection est d'obtenir une meilleure représentation de l'espace des images MS en procédant à un changement d'espace. Cette représentation vise à réduire la redondance d'informations en décorrélant les images MS. Une des composantes a la particularité d'isoler les structures des images multispectrales du reste des informations relatives à la couleur. L'hypothèse de ce type de méthodes est d'admettre que ces structures sont similaires à celles présentes dans la modalité A . La substitution consiste à transformer la composante structure des images MS en une fonction de l'image A (Shettigara 1992).

5.2.1 Rappel sur la notion d'espace de représentation des couleurs et principe de ce type de méthodes

Une combinaison colorée de trois canaux permet de retrouver l'information de couleur. Généralement, les espaces de représentation des couleurs les plus courants utilisent les trois modalités correspondant aux couleurs primaires du spectre (R, V, B). Il est toutefois possible de considérer une composition colorée en fausses couleurs (PIR, R, V) ou tout autre combinaison (addition, division, soustraction...) de trois modalités, y compris des images hétérogènes telles que des données radars avec des données optiques.

Une représentation simple des couleurs est l'espace HSV (*hue saturation value*) pour teinte (H) saturation (S) et valeur (V). Le logiciel ENVI propose une fonction de fusion exploitant cet espace. HSV est une représentation en trois dimensions de l'espace de couleurs défini par Smith en 1978. La figure 5.1 schématise le cône de définition du modèle :

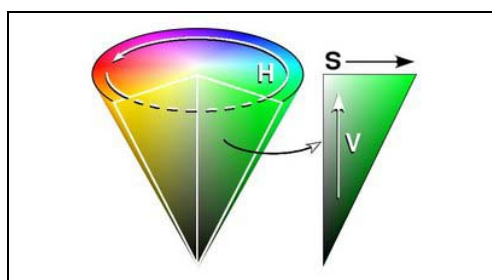


Figure 5.1 : le modèle de couleurs HSV de Smith.

Le haut du cône correspond à $V = 1$ et contient la couleur de valeur maximum correspondant au blanc. Le point à l'apex correspond au noir et possède la coordonnée $V = 0$. Les couleurs complémentaires sont symétriques par rapport à l'axe V et le rouge est fixé à l'angle $H = 0^\circ$. La valeur de S est un ratio s'étalant de 0 sur l'axe V jusqu'à 1 lorsque l'on se situe sur la surface extérieure du cône. Par exemple, d'après le graphique, on voit que la couleur jaune a pour coordonnées $H = \pi/4$, $S = 1$ et V prend une valeur quelconque entre 0 et 1 selon l'intensité du jaune. H et S sont donnés par des relations non linéaires de R , G et B . Par contre, la composante V est donnée par $V = \max(R, G, B)$, et ignore ainsi deux des modalités initiales.

Quelle que soit la combinaison choisie, les différentes composantes doivent être alignées géométriquement. Notons aussi l'importance de la dynamique des images sur la représentation des

couleurs. Si les dynamiques sont similaires, cela permet de ne pas privilégier une des modalités par rapport aux autres lors de la transformation.

Le problème posé aux développeurs de méthodes de type *COS* est de trouver un espace de représentation des couleurs adapté à l'enjeu de la fusion d'images satellites à haute résolution spatiale. Cette nouvelle représentation des modalités doit permettre d'isoler le contenu spectral (couleurs) des structures (fréquences spatiales) de l'image, et d'intervenir exclusivement sur cette dernière composante afin de limiter la distorsion radiométrique. Les modalités multispectrales d'origine sont rééchantillonnées afin d'adapter la taille des images à la modalité *A*. Le processus de fusion *COS* fonctionne de la manière suivante :

1. tout d'abord, une transformation directe (ou projection) est appliquée aux *N* modalités MS suréchantillonnées $(B_{kl})^{interp}_0$ avec, pour objectif, la transformation vers un espace de représentation dont l'une des composantes contient la plupart des structures présentes dans l'ensemble *B*,
2. la fusion intervient grâce à une substitution totale ou partielle de la composante des structures significatives par la modalité *A*. Par 'partielle', on entend toute combinaison, linéaire ou non, entre la composante à substituer et l'ensemble *A*, afin de pouvoir pondérer, et donc limiter, l'introduction d'informations issues de *A*,
3. et enfin, le retour à l'espace initial des modalités *B* s'effectue en appliquant la projection inverse de celle utilisée en 1.

La partie suivante présente la transformation *IHS* qui est présente dans la plupart des logiciels de traitement d'images qui disposent de fonctionnalités de fusion.

5.2.2 La méthode *IHS* linéaire

Il s'agit d'une méthode ancienne et très largement utilisée (Haydn *et al.* 1982, Carper *et al.* 1990, Chavez *et al.* 1991, Ehlers 1991, Pratt 1991, Tu *et al.* 2001) Une fois encore, seules trois bandes multispectrales peuvent être traitées simultanément. Après rééchantillonnage des modalités *R*, *V* et *B*, les coordonnées cartésiennes du modèle linéaire *IHS* définies par Pratt (1991) sont :

$$\left. \begin{aligned} \begin{pmatrix} I \\ v_1 \\ v_2 \end{pmatrix} &= \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} (R_1)^{interp}_0 \\ (G_1)^{interp}_0 \\ (B_1)^{interp}_0 \end{pmatrix} \\ H &= \tan^{-1}(v_2/v_1) \text{ avec } v_1 \neq 0 \\ S &= \sqrt{v_1^2 + v_2^2} \end{aligned} \right\} \text{ IHS} \quad \text{Eq. 5.2}$$

L'intensité est donc une combinaison linéaire des modalités placées en entrée, et correspond à la modalité qui va être remplacée dans le processus de fusion puisque, par construction, elle réunit les structures géométriques des modalités multispectrales. La nouvelle intensité *I'* est une fonction de la modalité *A* et de l'intensité *I*. L'insertion de la modalité Pan est soit totale, soit partielle, comme le montre respectivement les deux expressions de *I'* suivantes :

$$I' = A_0 \quad \text{Eq. 5.3}$$

$$I' = \alpha I + (1-\alpha) A_0 \quad \text{Eq. 5.4}$$

avec $\alpha \in [0, 1]$. La dynamique de *A* est souvent ajustée à celle de *I*.

La transformée inverse s'applique de la façon suivante :

$$\left. \begin{aligned} v_1 &= S \cos(H) \\ v_2 &= S \sin(H) \\ \begin{pmatrix} R_0^* \\ G_0^* \\ B_0^* \end{pmatrix} &= \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -2/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix} \begin{pmatrix} I' \\ v_1 \\ v_2 \end{pmatrix} \end{aligned} \right\} \text{ IHS}^{-1} \quad \text{Eq. 5.5}$$

Tu *et al.* (2001) ont proposé une autre formulation de la transformée *IHS* linéaire. Soit I' la nouvelle intensité à haute résolution spatiale. Le remaniement des équations 5.2 et 5.5 aboutit à la nouvelle modalité B_{k0}^* :

$$B_{k0}^* = (B_{kl})_0^{interp} + \delta, \text{ avec } \delta = I' - I \quad \text{Eq. 5.6}$$

Des versions plus élaborées et non linéaires de la transformée *IHS* peuvent être consultées dans Wald (2000) et Gonzalez et Woods (1992).

5.2.3 La méthode ACP

L'Analyse en Composantes Principales (ACP), connue aussi sous le nom de la technique de Karhunen-Loeve, est un outil mathématique qui peut s'appliquer sur un nombre N quelconque de modalités (Chavez 1991). Par rapport à *IHS*, cette transformée est donc particulièrement bien appropriée en imagerie multi- ou hyperspectrale. Elle applique une transformation linéaire de l'espace multispectral vers l'espace des vecteurs propres des données considérées. Les N composantes ACP obtenues sont orthogonales, ce qui signifie qu'aucune composante n'est linéairement corrélée avec une autre. La variance totale des N images originales est répartie entre les N composantes ACP de telle sorte que la première composante possède la plus grosse part de cette variance, puis que la variance décroît en passant d'une composante à la suivante.

Notons que l'utilisation de la transformée ACP peut constituer la première étape pour une transformation de l'espace des N modalités MS initiales ($N > 3$) vers un espace des couleurs nécessitant seulement trois modalités. En effet, en sélectionnant les trois premières composantes, l'ACP permet de réduire la quantité d'information et d'en extraire les trois canaux les plus représentatifs en terme de variance de ces images.

Soit V la matrice unitaire dont les colonnes représentent les vecteurs propres $V = (v_1, \dots, v_n)$, avec $v_k = (v_{1,k}, \dots, v_{n,k})^t$. La $k^{\text{ième}}$ composante P_k de la transformée ACP est donnée par :

$$\begin{pmatrix} \dots \\ \vec{P}_k \\ \dots \end{pmatrix} = \begin{pmatrix} v_{1,1} & \dots & \dots & v_{1,N} \\ \dots & \dots & \dots & \dots \\ v_{k,1} & \dots & \dots & v_{k,N} \\ \dots & \dots & \dots & \dots \\ v_{N,1} & \dots & \dots & v_{N,N} \end{pmatrix} \begin{pmatrix} (B_{1,1})_0^{interp} \\ \dots \\ (B_{N,1})_0^{interp} \end{pmatrix} \quad \text{Eq. 5.7}$$

donc :

$$P_k = \sum_{p=1}^N v_{p,k} (B_{p,1})_0^{interp} \quad \text{Eq. 5.8}$$

Si C correspond à la matrice de covariance (cov) de l'ensemble B où $C(i, j) = cov(B_i, B_j)$, alors C est une matrice symétrique, ce qui implique aussi que cette matrice est diagonalisable :

$$V^t C V = \begin{pmatrix} \delta_1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \delta_n \end{pmatrix} \quad \text{Eq. 5.9}$$

où $\{\delta_1, \dots, \delta_n\}$ sont les valeurs propres rangées par ordre décroissant : $\delta_1 > \dots > \delta_n$. δ_k correspond à la variance totale contenue dans la $k^{\text{ième}}$ composante. La variance totale est égale à la somme des δ_k avec $k = \{1..n\}$. La modalité fusionnée B_{k0}^* s'obtient en plaçant $P_1 = A_0$ et par la relation $B_{k0}^* = V^{-1} P_k$.

Une autre possibilité est l'utilisation de la matrice de corrélation au lieu de la matrice de covariance. Celle-ci a l'avantage de procéder à une normalisation des données et permet de traiter des signaux dont les dynamiques sont très différentes, sans en favoriser un plutôt qu'un autre, mais malheureusement elle introduit une distorsion radiométrique importante.

5.3 Les méthodes de type contribution spectrale relative

5.3.1 Hypothèse de base de ces méthodes

Une hypothèse fondamentale régit le fonctionnement des méthodes de fusion de type contribution spectrale relative : la modalité Pan à basse résolution spatiale est équivalente à une combinaison linéaire des modalités MS :

$$A_I = \sum_{k=1}^N \alpha_k B_{k1} \quad \text{avec } k = \{1..N\} \quad \text{Eq. 5.10}$$

où la valeur des pondérations des bandes α_k dépend de la méthode de fusion considérée. Cette pondération tient généralement compte de la proportion de la réponse spectrale de la $k^{\text{ième}}$ modalité MS par rapport à celle de la modalité Pan. L'objectif de ces méthodes est de multiplier chaque modalité MS suréchantillonnée par une fonction des hautes fréquences de la modalité Pan. Nous regroupons ici toutes les méthodes qui ne font pas partie des approches multiéchelles ou multirésolution.

5.3.2 La méthode Brovey

La méthode de fusion Brovey (Gillespie *et al.* 1987, Zhou 1998, Liu *et al.* 2000) est la plus simple puisque les coefficients α_k sont fixés à 1. L'image fusionnée B_{k0}^* s'obtient par un produit pixel à pixel entre la modalité correspondante à basse résolution rééchantillonnée $(B_{k1})^{interp}_0$ et un terme fonction de la modalité Pan haute résolution, divisé par la somme des modalités MS suréchantillonnées :

$$B_{k0}^* = (B_{k1})^{interp}_0 \frac{A_0}{\sum_{k=1}^N (B_{k1})^{interp}_0} \quad \text{Eq. 5.11}$$

Les moyennes et les dynamiques des images ne sont ajustées, ce qui entraîne une importante distorsion radiométrique, c'est-à-dire un non-respect de la propriété de cohérence. Les prochaines versions de cette méthode ajustent la moyenne de la modalité fusionnée à celle de cette même modalité avant la fusion.

5.3.3 La méthode CN (Color Normalized)

Cette méthode constitue une amélioration de la méthode Brovey. Pour $N = 3$ modalités MS, elle s'écrit :

$$B_{k0}^* = \frac{3((B_{k1})^{interp}_0 + 1)(A_0 + 1)}{3 + \sum_{k=1}^N (B_{k1})^{interp}_0} - 1 \quad \text{Eq. 5.12}$$

Les valeurs sont en niveaux de gris. Les dynamiques de toutes les images de l'expression doivent être similaires. Les constantes de faibles valeurs permettent de s'affranchir des perturbations liées aux divisions par 0.

5.3.4 La méthode P+XS (CNES)

Cette méthode a été développée par le CNES, l'agence spatiale française, pour le système SPOT pour produire des images multispectrales avec une résolution de 10 m. L'hypothèse est que la demi-somme des luminances des bandes XS1 et XS2 est égale à la luminance du Pan. Une critique de cette hypothèse a été faite par Mangolini (1994). Cette méthode peut s'appliquer à tout ensemble d'images vérifiant cette hypothèse.

Soient L_{A0} , $L_{XS1,0}$ et $L_{XS2,0}$ les luminances d'un pixel à 10 m de résolution respectivement dans les images A , $XS1^*$ et $XS2^*$, ces deux dernières correspondant aux modalités SPOT fusionnées à haute résolution spatiale. L'hypothèse s'écrit alors :

$$L_{A0} = \frac{1}{2} (L_{XS1,0}^* + L_{XS2,0}^*) \quad \text{Eq. 5.13}$$

Une seconde hypothèse est nécessaire pour décrire la distribution de la luminance L_{A0} entre $L_{XS1,0}$ et $L_{XS2,0}$. La méthode $P+XS$ suppose que le ratio entre les luminances dans les deux bandes est constant avec le changement de $res1$ vers $res0$:

$$L_{XS1,0}^* / L_{XS2,0}^* = L_{XS1,1} / L_{XS2,1} \quad \text{Eq. 5.14}$$

Cette hypothèse n'est vraie que si le pixel représentant un carré de 20 m par 20 m au sol est composé exactement de quatre pixels de 10 m par 10 m ayant le même comportement spectral. La manipulation des équations aboutit aux expressions :

$$L_{XS1,0}^* = 2 L_{A0} (L_{XS1,1})^{interp_0} / ((L_{XS1,1})^{interp_0} + (L_{XS2,1})^{interp_0}) \quad \text{Eq. 5.15}$$

$$L_{XS2,0}^* = 2 L_{A0} ((L_{XS2,1})^{interp_0} / ((L_{XS1,1})^{interp_0} + (L_{XS2,1})^{interp_0})) \quad \text{Eq. 5.16}$$

Ces formules produisent des images synthétiques $XS1_0^*$ et $XS2_0^*$. A aucun moment la troisième modalité MS ($XS3$) de SPOT n'intervient dans les équations 5.11 à 5.14. Les auteurs synthétisent l'image $XS3_0^*$ par une simple duplication des pixels de la bande $XS3_1$.

On notera que contrairement aux précédentes qui s'appliquent aux niveaux de gris, la méthode $P+XS$ exploite des luminances qui résultent de l'application d'une fonction d'étalonnage aux niveaux de gris.

5.4 Le concept ARSIS

5.4.1 Présentation du concept

ARSIS est l'acronyme de « Amélioration de la Résolution Spatiale par Injection de Structures ». Le concept *ARSIS* a été développé de manière très générale, en ne faisant référence à aucun outil mathématique utilisé pour ses déclinaisons informatiques (Ranchin et Wald 2000). Le concept *ARSIS* suppose que l'information manquante de l'image multispectrale de basse résolution B_{kl} est composée uniquement des informations de hautes fréquences spatiales. Il considère que ces informations peuvent être synthétisées à partir de l'ensemble des données disponibles à toutes les échelles et dans toutes les modalités. Il s'appuie sur les techniques multiéchelles (ou modèles multiéchelles, soit en anglais *Multi-Scale Models* ou *MSM*), pour modéliser puis injecter les hautes fréquences manquantes. Ces techniques font appel à des outils mathématiques tels que les espaces vectoriels et le filtrage numérique, qui permettent d'obtenir une description hiérarchique des structures spatiales d'une image (Ranchin et Wald 1992).

En traitement d'image, les décompositions hiérarchiques ont de manière générale une représentation pyramidale (figure 5.2). Les différents étages de la pyramide correspondent à des versions de plus en plus grossières de l'image originale, c'est-à-dire qu'à chaque étage, les hautes fréquences spatiales de l'étage sous-jacent ont été filtrées. Ce sont les images de contexte ou d'approximation qui, d'après les notations proposées en introduction, s'écriront $A_0, A_1, A_2... A_n$ pour la pyramide A , et $B_{k0}, B_{k1}, B_{k2}... B_{kn}$ pour la pyramide B_k . La différence d'information entre deux plans d'approximation successifs correspond aux détails manquants dans l'approximation la plus grossière.

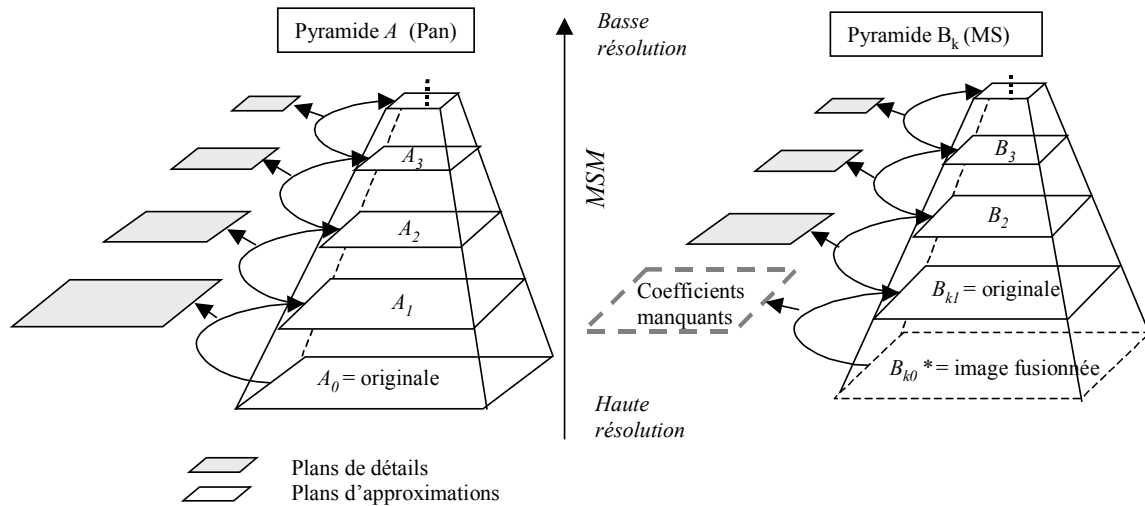


Figure 5.2 : décomposition hiérarchique de l'information.

L'enjeu du concept *ARSIS* est d'injecter de l'information pour compléter le plan en pointillés correspondant aux détails manquants dans la pyramide MS de la figure 5.2. Sauf cas particuliers, les détails ne peuvent directement être égaux à ceux de l'image *A* car l'image synthétisée B_{k0}^* ne serait pas équivalente à « ce que verrait le capteur B_k s'il possédait la résolution spatiale du capteur *A* ». Une transformation doit être appliquée pour convertir les informations fournies par la représentation multiéchelle de l'image *A* en celles adaptées à la synthèse de B_{k0}^* . Il faut donc établir une relation qui lie le contenu spatial de ces deux ensembles aux échelles spatiales où l'information est présente dans les deux représentations pour pouvoir adapter et transformer l'information à injecter.

Par souci de simplicité, nous nous plaçons dans le cas où le *MSM* permet directement de passer de $res0$ à $res1$ puis à $res2$ etc... Cette situation n'est pas celle créée par l'utilisation d'un algorithme dyadique pour le ratio 4 car dans ce cas, le *MSM* passe de la résolution $res0$ à $2res0$, de $2res0$ à $res1$ puis de $res1$ à $2res1$... ; les résolutions $2res0$, $2res1$ etc... correspondent alors à des résolutions spatiales intermédiaires qui ne figurent pas dans les notations.

La figure 5.3 présente une vision globale du concept *ARSIS*. Les quatre opérations y sont représentées par application successive de trois classes de modèles : classe 1 : *MSM*, classe 2 : *IMM*, classe 3 : *HRIMM*, et enfin une ultime étape de reconstruction utilisant l'inverse du modèle *MSM* : MSM^{-1} :

- ♦ les *MSM* produisent une description hiérarchique de l'information des structures spatiales de l'image. Par exemple, dans la littérature, les auteurs font appel aux modèles multiéchelles basés sur la transformée en ondelettes, la pyramide Laplacienne généralisée (*GLP* pour *Generalized Laplacian Pyramid*, Aiazzi et al. 1997), ou la dérivée seconde de la fenêtre d'apodisation. Appliqué à une image, le *MSM* fournit une (ou plusieurs) image de détails qui représente les hautes fréquences spatiales, et une image d'approximation qui regroupe les basses fréquences de l'image. La première itération du *MSM* sur la modalité *A* donne donc une ou plusieurs images de structures comprises entre $res0$ et $res1$ et une image de structures supérieures ou égales à $res1$ (approximation 2). La variabilité spatiale d'une image peut alors être modélisée. Le *MSM* doit être inversible (MSM^{-1}) afin de réaliser la synthèse des informations de hautes fréquences,
- ♦ la classe de modèles inter-modalités (en anglais *Inter-Modality Model*, soit *IMM*) s'intéresse à la transformation des structures spatiales avec le changement de modalités. L'*IMM* modélise les relations entre les détails et/ou les approximations observés dans l'image *A* et ceux observés dans l'image B_k . Cette relation peut suivre un modèle linéaire, affine, ou être d'une forme mathématique plus complexe,
- ♦ et enfin, la classe des modèles inter-modalités à haute résolution spatiale (en anglais *High Resolution IMM*, soit *HRIMM*) qui doit synthétiser le plan de détails manquants. Le *HRIMM*

résulte de l'ensemble des informations à notre disposition : des approximations et détails à basse résolution, des paramètres de l'*IMM* et du plan de détails haute fréquence existant dans la pyramide *A*.

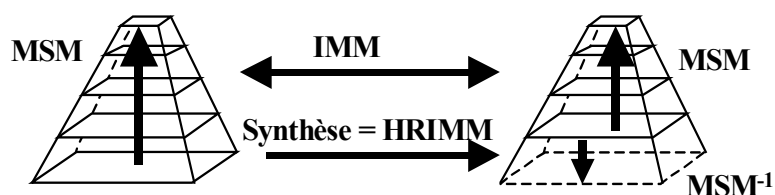


Figure 5.3 : les trois classes de modèles du concept *ARSIS*.

La figure 5.4 détaille les classes de modèles et leurs interactions avec les différents plans disponibles. Les entrées du procédé de fusion sont la modalité *A* à *res0*, et la modalité *B_k* à basse résolution *res1*. Les opérations sont effectuées comme suit :

1. tout d'abord, le *MSM* est utilisé pour décomposer les images *A* et *B_k* originales et en calculer les détails et approximations successives,
2. les approximations et les détails connus des deux pyramides pour chaque résolution sont utilisés pour ajuster les paramètres de l'*IMM*,
3. à partir de l'*IMM*, on définit le *HRIMM* qui convertit les détails connus de *A* pour en déduire ceux de l'image *B_{k1}* qui étaient inconnus à *res1*,
4. enfin, l'image *B_{k0}*^{*} est synthétisée par inversion du *MSM* (*MSM*⁻¹) de *res1* vers *res0*.

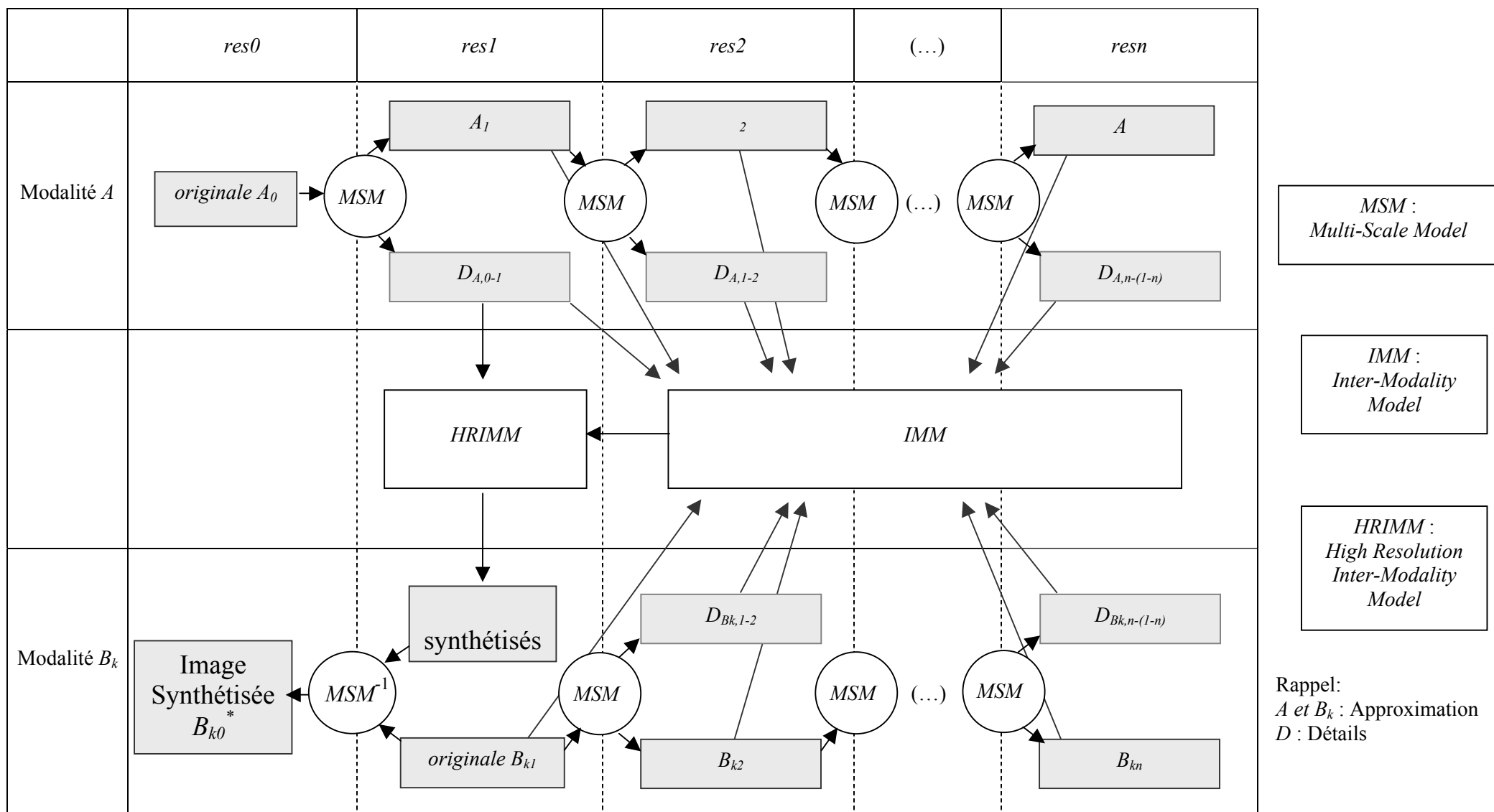


Figure 5.4 : schéma descriptif du concept ARSIS (d'après Wald 2000).

Nous allons décrire les fonctions mathématiques des *IMM* et *HRIMM*. Nous ne développons pas ici celles qui définissent les *MSM* car elles sont largement décrites dans la littérature.

Soient D_A et D_{B_k} les plans de détails respectivement pour les pyramides de A et de B_k . Le plan de détails $D_{A,0-1}$ représente les informations présentes entre les deux plans d'approximation successifs $res0$ et $res1$ pour la pyramide A . Pour simplifier l'écriture, nous nous plaçons dans le cas où l'analyse multiéchelle délivre un unique plan de détails par résolution. Tel est le cas par exemple de la transformée *ATWT* (*A Trou Wavelet Transform*) qui est non directionnelle et non décimée (Dutilleux 1989) ou de la pyramide Laplacienne généralisée *GLP* (*Generalized Laplacian Pyramid*) (Aiazzi et al. 2002). L'écriture générale des plans de détails successifs pour les deux pyramides :

$$D_A = \{D_{A,0-1}, D_{A,1-2}, D_{A,2-3} \dots\} \text{ et } D_{B_k} = \{D_{B_k,0-1}, D_{B_k,1-2}, D_{B_k,2-3} \dots\} \quad \text{Eq. 5.17}$$

Dans d'autres cas, comme pour les ondelettes directionnelles de l'analyse multirésolution de *Mallat* (1989), le plan d'ondelettes n'est plus unique, et doit à son tour être indicé par D , V et H pour diagonal, vertical et horizontal. Nous recommandons l'utilisation de filtres symétriques

Si on se réfère à la figure 5.4 qui décrit le concept *ARSIS*, l'*IMM* doit estimer des paramètres pour modéliser les relations entre les images de A et B_k aux échelles où les informations sont présentes simultanément dans les deux pyramides. Soit f_1 la fonction mathématique de l'*IMM*, alors f_1 peut s'écrire :

$$D_{B_k,1-2} = f_1(D_{A,1-2}, D_{A,2-3}, \dots, D_{B_k,1-2}, D_{B_k,2-3}, \dots, A_1, A_2, A_3, \dots, B_{k1}, B_{k2}, B_{k3}, \dots) \quad \text{Eq. 5.18}$$

Une autre possibilité est une fonction basée sur l'établissement d'une relation liant l'image d'approximation B_{k2} et les autres plans présents simultanément dans les deux pyramides de données :

$$B_{k2} = f_1(D_{A,1-2}, D_{A,2-3}, \dots, D_{B_k,1-2}, D_{B_k,2-3}, \dots, A_1, A_2, A_3, \dots, B_{k1}, B_{k2}, B_{k3}, \dots) \quad \text{Eq. 5.19}$$

Soit f_2 la fonction mathématique qui décrit le *HRIMM*. Cette fonction s'écrit à son tour comme une fonction des plans de détails et approximations des deux pyramides, ainsi que le plan de détails à haute résolution spatiale exclusivement présente dans la pyramide Pan $D_{A,0-1}$; alors :

$$D_{B_k,0-1} = f_2(D_{A,0-1}, D_{A,1-2}, D_{A,2-3} \dots, D_{B_k,1-2}, D_{B_k,2-3}, \dots, A_1, A_2, A_3, \dots, B_{k1}, B_{k2}, B_{k3}, \dots) \quad \text{Eq. 5.20}$$

A notre connaissance, peu d'auteurs ont apporté une attention particulière au *HRIMM*. La plupart du temps, f_2 est définie identique à f_1 . On assiste rarement dans la littérature à un ajustement des paramètres issus de l'*IMM* vers le *HRIMM*.

Notons que les auteurs du concept *ARSIS* préconisent l'usage de luminances en entrée du processus de fusion afin de mieux prendre en compte la physique de l'observation. Cela n'est pas toujours nécessaire.

5.4.2 Quelques exemples de réalisations du concept

Le concept *ARSIS* a été décliné en plusieurs implantations informatiques. Une méthode de fusion, ou implantation, de type *ARSIS* correspond à une association entre un modèle de chaque classe : un *MSM*, un *IMM* et un *HRIMM*. Concernant les *MSM*, nous avons utilisé les outils *ATWT*, *Mallat* et *GLP*. Pour pouvoir tester simplement différentes stratégies ou implantations, nous avons développé un atelier logiciel en IDL qui comprend également plusieurs *IMM* et *HRIMM*. L'interface graphique de cet atelier fusion ajoutée à celle pour l'évaluation de la qualité a été placée en annexe (annexe 1). Il est possible d'associer dans cet atelier un modèle de chaque classe, soit trois modèles, pour créer une implantation (Thomas et al. 2005). Nous présentons ici après quelques exemples de couples *IMM-HRIMM* que nous associons généralement à *ATWT*.

5.4.2.1 La méthode de Pradines (1986)

Dans le cas de la méthode de Pradines (1986), l'extraction des hautes fréquences de la modalité Pan est effectuée par division entre la modalité Pan et une version basse résolution de cette dernière, et peut ainsi être assimilée à un *MSM*. L'équation permettant d'obtenir la modalité fusionnée B_{k0}^* est :

$$B_{k0}^* = (B_{kl})^{interp}_0 A_0 / A_l \quad \text{Eq. 5.21}$$

Le filtre utilisé pour synthétiser l'approximation A_l est un Heaviside, c'est-à-dire une fonction échelon, dont la réponse spectrale (le sinus cardinal) engendre de nombreux rebonds. Des améliorations de cette méthode concernant la synthèse de la modalité A_l sont la méthode *Pixel Block Intensity Modulation* de Liu et Moore (1998) et la méthode *Smoothing Filter-based Intensity Modulation* de Liu (2000).

5.4.2.2 La méthode de Price (1999)

Price (1999) suit aussi implicitement cette procédure. L'extraction des hautes fréquences est produite non pas par division, mais par différence entre la modalité Pan et sa version plus grossière. Il s'agit du principe repris par Wang *et al.* (2005) présenté en début de ce chapitre :

$$B_{k0}^* = (B_{kl})^{interp}_0 + \alpha_{A_l, B_{kl}} (A_0 - (A_l)^{interp}_0) \quad \text{Eq. 5.22}$$

où le coefficient $\alpha_{A_l, B_{kl}}$ est calculé sur une fenêtre glissante, typiquement de taille 3 par 3 ou 5 par 5 pixels. Ce coefficient résulte de la régression linéaire entre les images suivantes :

$$B_{kl} = \alpha_{A_l, B_{kl}} A_l + b \quad \text{Eq. 5.23}$$

L'avantage de cette dernière méthode est d'offrir la possibilité de capturer les inversions de contraste grâce au coefficient $\alpha_{A_l, B_{kl}}$ qui peut être négatif en fonction de la corrélation entre les deux images B_{kl} et A_l . Une autre méthode très proche de celle-ci est la méthode appelée *Local Correlation Modeling* de Diemer et Hill (2000).

5.4.2.3 Couple M1

Ce couple ne procède pas à une estimation de paramètres (absence d'*IMM*) (Ranchin et Wald 2000). *M1* se résume à un *HRIMM* simple correspondant à la fonction identité. L'image MS à *res0* est donc directement synthétisée en introduisant sans pré-traitement le premier plan de détails disponible dans la pyramide A , $D_{A,0-1}$:

$$D_{B_{k0},0-1} = D_{A,0-1} \quad \text{Eq. 5.24}$$

La figure 5.5 schématise l'implantation de ce couple :

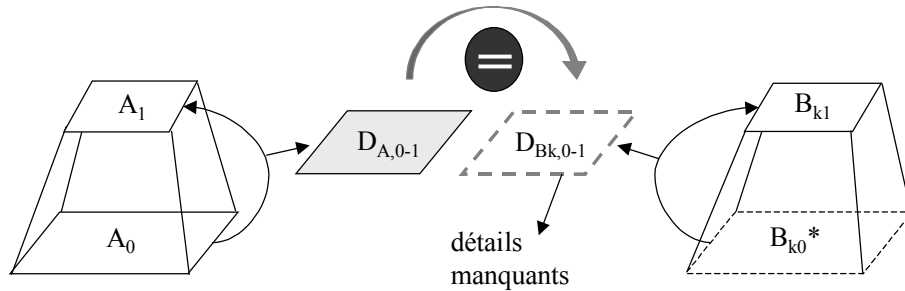


Figure 5.5 : schéma récapitulatif du couple *IMM-HRIMM M1*.

Notre laboratoire utilise *M1* en combinaison avec un *MSM* de type transformée en ondelettes. D'autres méthodes connues suivent cette méthodologie, comme par exemple la méthode de Chavez *et al.* (1991) notée *HPF* pour *High Pass Filter*. Cette méthode de fusion procède à l'extraction des hautes fréquences spatiales de la modalité Pan par convolution avec un filtre passe-haut Laplacien de taille 3 par 3 pixels. La différence essentielle entre ces implantations se situe dans la réponse spectrale des filtres utilisés, dont certains présentent des phénomènes de rebonds ou de repliement spectral.

5.4.2.4 Couple M2

Les deux couples suivants *M2* et *M3* reposent sur la même hypothèse : pour une modalité donnée, il existe une forte relation affine entre les détails de l'image A et ceux de B pour des échelles supérieures à *res1* (Ranchin et Wald 2000). On admet que cette relation reste inchangée avec le

changement d'échelle $res1-res0$ ($IMM = HRIMM$). Cette hypothèse est moins exigeante que celle qui impose une forte corrélation entre A et B_k pour les méthodes de projection et de substitution.

$M2$ cherche à définir une relation de type affine, basée sur l'ajustement des moyennes et des variances des premiers plans de détails disponibles simultanément dans les deux pyramides ($D_{A,1-2}$ et $D_{Bk,1-2}$). L'histogramme du plan de détail $D_{A,1-2}$ est adapté à celui de $D_{Bk,1-2}$. L' IMM se définit de la manière suivante :

$$D_{Bk,1-2} = f_2(D_{A,1-2}) = aD_{A,1-2} + b \quad \text{Eq. 5.25}$$

avec a et b :

$$a = \sqrt{\frac{v(D_{Bk,1-2})}{v(D_{A,1-2})}} \quad \text{Eq. 5.26}$$

$$b = m(D_{Bk,1-2}) - a m(D_{A,1-2}) \quad \text{Eq. 5.27}$$

où m et v définissent les opérateurs de moyenne et de variance. a et b seront différents pour chaque modalité. Pour le $HRIMM$, les détails manquants sont alors déterminés par :

$$D_{Bk,0-1} = f_1(D_{A,0-1}) = aD_{A,0-1} + b \quad \text{Eq. 5.28}$$

et donc ici $f_1 = f_2$, c'est-à-dire que $IMM = HRIMM$. La figure 5.6 explique le principe des couples $M2$ et $M3$.

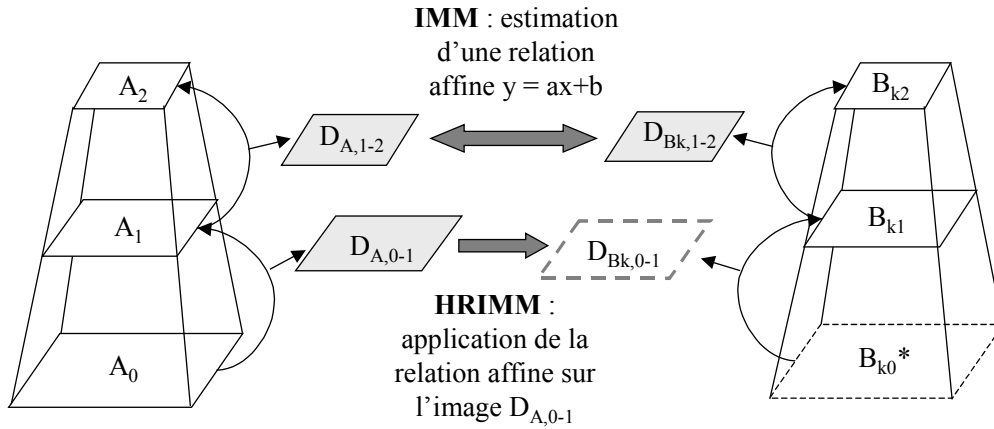


Figure 5.6 : schéma récapitulatif des deux couples $M2$ et $M3$.

5.4.2.5 Couple $M3$

Cette association fonctionne de manière similaire au couple précédent, mais le coefficient a est obtenu soit par minimisation de l'axe d'inertie, soit par les moindres carrés. Dans le cas des moindres carrés, si cov correspond à l'opérateur de covariance et var à celui de la variance, l'écriture de a est :

$$a = cov(D_{Bk,1-2}) / var(D_{A,1-2}) \quad \text{Eq. 5.29}$$

Par contre, l'écriture pour la minimisation du centre d'inertie est un peu plus complexe. a se définit de la manière suivante :

$$a = \frac{var(D_{Bk,1-2}) - var(D_{A,1-2}) + \sqrt{(var(D_{Bk,1-2}) - var(D_{A,1-2}))^2 + 4cov^2(D_{Bk,1-2}, D_{A,1-2})}}{2cov(D_{Bk,1-2}, D_{A,1-2})} \quad \text{Eq. 5.30}$$

b restant identique à l'implantation précédente. Le choix de la méthode de minimisation est un paramètre d'entrée de notre plate-forme *fusion* en IDL.

5.4.2.6 Couple $M4$ ou RWM (Ranchin-Wald-Mangolini)

L'approche de cette dernière implantation conserve l'idée de relation affine. Cependant, jusqu'à maintenant les coefficients a et b étaient évalués de manière globale sur toute l'image. Ici, l'estimation de ces deux paramètres s'effectue localement, à l'aide d'une fenêtre glissante dont la taille dépend des résolutions spatiales des modalités initiales (Ranchin *et al.* 1994, Wald *et al.* 1997).

Ranchin *et al.* (2003) déterminent la taille t de la fenêtre glissante à partir des résolutions des deux ensembles initiaux A et B respectivement $res0$ et $res1$ par la relation suivante :

$$t = 7(2res1)/res0 + 1 \quad \text{Eq. 5.31}$$

Notons que cette formule est fautive (dixit ses auteurs). Nous proposons en annexe un récapitulatif des différentes tailles de fenêtre d'estimation par méthode de fusion (annexe 4).

Le recours à la fenêtre glissante dans l'évaluation aboutit à des images de coefficients a et b , au lieu d'un coefficient unique pour l'ensemble de l'image. Une étape supplémentaire vient conditionner l'application des paramètres a et b calculés pour un pixel donné ; la similarité entre les détails $D_{A,1-2}$ et $D_{Bk,1-2}$ est localement testée par le calcul des coefficients de corrélation cc locaux. L'application conditionnelle à haute résolution spatiale (*HRIMM*) de la relation affine dépendra de la valeur de la corrélation. Cet élément marque à la fois l'originalité et la supériorité de *RWM* par rapport aux modèles précédents. Si le *MSM* utilisé est *ATWT*, la taille des images est la même quelque soit la résolution spatiale et l'*IMM* reste identique au *HRIMM*, c'est-à-dire que $f1 = f2$. Si, par contre, on choisit d'utiliser un *MSM* décimé, il faudra suréchantillonner les images de a et de b , ce qui implique $f1 \neq f2$. Deux seuils $s1$ et $s2$ pour la valeur absolue du coefficient de corrélation ($|cc|$) vont rentrer en jeu dans la description du modèle. Les valeurs de ces seuils sont respectivement 0.01 et 0.7.

Quatre situations sont alors envisageables. Si $|cc| < s1$, alors on procède à un ajustement d'histogramme entre les deux fenêtres des images $D_{Bk,1-2}$ et $D_{A,1-2}$ (comme pour le modèle *M2*), que a prend aussi en compte le signe (*sign*) du produit local des détails :

$$a_{|cc|<s1} = (sign) \sigma(D_{Bk,1-2})/\sigma(D_{A,1-2}) \quad \text{Eq. 5.32}$$

où *sign* correspond au signe issu du produit entre $D_{A,1-2}(i, j)$ et $D_{Bk,1-2}(i, j)$ au pixel (i, j) considéré. Pour éviter les divisions par 0, si $\sigma(D_{A,1-2}) < 10^{-5}$, alors $a = 0$ ce qui fixe le coefficient d'ondelettes à cet endroit précis à 0, produisant ainsi un simple rééchantillonnage de l'image MS basse résolution (Ranchin *et al.* 1993).

La seconde situation possible est si $|cc| > s2$ pour un pixel donné. Alors $a_{|cc|>s2}$ se calcule en appliquant un ajustement des deux fenêtres par les moindres carrés ou l'axe d'inertie selon le choix de l'utilisateur (comme pour le modèle *M3*). Il faut prendre garde de s'affranchir des divisions par 0. Si l'ajustement choisi est l'axe d'inertie, on admet qu'on a une division par 0 dès lors que $cov(D_{Bk,1-2}, D_{A,1-2}) < 10^{-2}$, et dans le cas des moindres carrés, lorsque $var(D_{A,1-2}) < 10^{-2}$. Dans ces deux situations, on fixe $a_{|cc|>s2} = a_{|cc|<s1} = (sign) \sigma(D_{Bk,1-2})/\sigma(D_{A,1-2})$.

Si le coefficient de corrélation possède une valeur comprise entre $s1$ et $s2$ ($s1 < |cc| < s2$), on utilise une combinaison linéaire entre les deux valeurs de a calculées précédemment :

$$a = a_{|cc|<s1} + \left(\frac{|cc| - s2}{s1 - s2} \right) (a_{|cc|>s2} - a_{|cc|<s1}) \quad \text{Eq. 5.33}$$

Enfin, en cas de corrélation insuffisante, c'est-à-dire inférieure au seuil $s1$, on place a à 0 et donc un simple suréchantillonnage est effectué. Quelque soit la situation, l'ordonnée à l'origine b résulte de la formule :

$$b = m(D_{Bk,1-2}) - a m(D_{A,1-2}) \quad \text{Eq. 5.34}$$

Le schéma 5.7 synthétise l'ensemble de ces étapes.

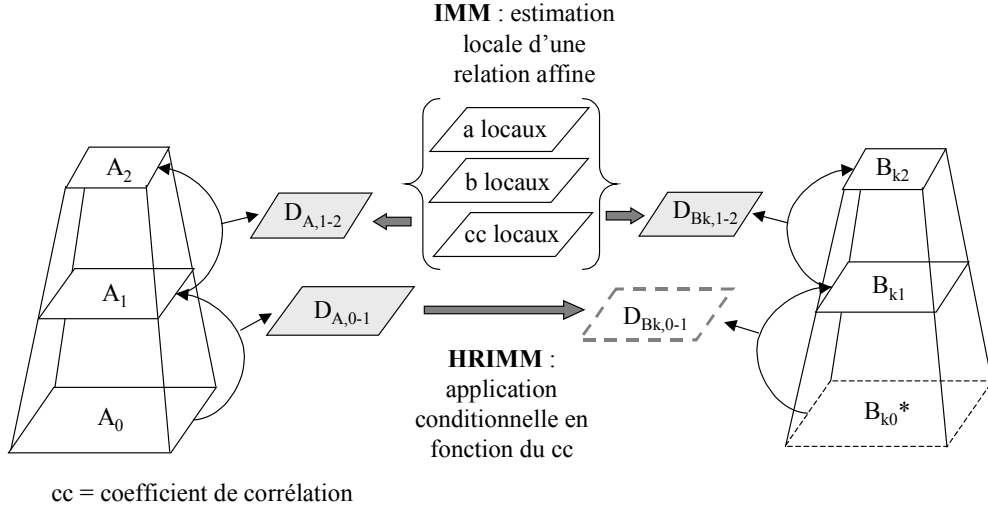


Figure 5.7 : schéma récapitulatif du couple *RWM*.

5.4.2.7 Couple *CBD* pour « Context-Based Decision »

Aiazzi *et al.* (2001) ont défini un couple *IMM-HRIMM CBD* pour désigner une approche basée contexte. Notons que cette méthode de fusion peut aussi être trouvée dans la littérature sous le nom de *AABP*, du nom de ces auteurs : Aiazzi, Alparone, Baronti et Pippi. Ces auteurs utilisent le modèle multiéchelle *GLP*. Cette association est relativement proche de la précédente *RWM* car les coefficients a et b qui lient les plans de détails sont eux aussi calculés sur des fenêtres glissantes. Leur application à haute résolution est aussi conditionnée par la valeur absolue d'un coefficient de corrélation. Par contre, cette fois-ci, la relation est linéaire (le coefficient b est implicitement placé à 0) et a et cc ne sont plus évalués à partir des plans de détails, mais à partir des plans de contexte. Mathématiquement, l'*IMM* est le produit pixel à pixel suivant :

$$B_{k,1-2} = f_1(A_2) = a A_2 \quad \text{Eq. 5.35}$$

La fonction *HRIMM* s'applique cependant toujours sur les plans de coefficients de détails :

$$D_{Bk,0-1} = f_2(D_{A,0-1}) = a D_{A,0-1} \quad \text{Eq. 5.36}$$

Le *HRIMM* est donc différent de l'*IMM* puisque ces deux fonctions s'appliquent sur des objets différents.

D'après les auteurs, les coefficients a et cc sont calculés sur une fenêtre glissante dont la taille dépend du satellite dont proviennent les images. Dans le cas de SPOT1 à SPOT4, la fenêtre a une taille 7 par 7 pixels, et pour Ikonos, de 9 par 9 pour le *MSM GLP*.

Soient $\sigma(B_{kl})$ et $\sigma(A_l)$ les écart-types des approximations B_k et A à res_l . Si la valeur absolue du coefficient de corrélation ($|cc|$) entre les contextes A_l et B_{kl} est supérieure au seuil s , avec s pouvant prendre une valeur située dans l'intervalle $[0.3, 0.6]$, alors a est donné par :

$$a = \min(\sigma(B_{kl}) / (\sigma(A_l) + 1), 3) \quad \text{Eq. 5.37}$$

et dans le cas où $|cc| < s$, a est fixé à 0.

5.4.2.8 Couple *SDM* pour « Spectral Distortion Minimizing »

La méthode de fusion développée par Garzelli *et al.* (2004) suit une approche vectorielle car toutes les modalités MS sont synthétisées de manière simultanée. Pour un pixel (i, j) donné et à la résolution res_0 , on rappelle qu'un vecteur multispectral s'écrit :

$$B_0^*(i, j) = \{B_{k0}^*(i, j) \dots, k = 1..n\} \quad \text{Eq. 5.38}$$

De la même façon, $(B_l)^{interp}_0(i, j)$ est le vecteur multispectral dont chaque coordonnée est donnée par la valeur du pixel (i, j) provenant de chaque modalité MS originale suréchantillonnée pour atteindre les dimensions de la modalité A_0 :

$$(B_l)^{interp}_0(i, j) = \{(B_{kl})^{interp}_0(i, j) \dots, k = 1..n\} \quad \text{Eq. 5.39}$$

Soit enfin $D_{Bk,0-1}(i, j)$ le vecteur de détails à injecter. Ces trois vecteurs sont reliés par la relation suivante :

$$B_0^*(i, j) = (B_1)^{interp}_0(i, j) + D_{Bk,0-1}(i, j) \quad \text{Eq. 5.40}$$

rejoignant ainsi la formulation du modèle *GIF* de Wang (2005) décrit en introduction de ce chapitre.

Le vecteur de détails à synthétiser est lié à celui du Pan par la relation :

$$D_{Bk,0-1}(i, j) = \alpha_k(i, j) D_{A,0-1}(i, j) \quad \text{Eq. 5.41}$$

où le produit entre le vecteur d'images α et l'image de détails du Pan est un produit pixel à pixel. Alors,

$$B_0^*(i, j) = (B_1)^{interp}_0(i, j) + \alpha_k(i, j) D_{A,0-1}(i, j) \quad \text{Eq. 5.42}$$

La $k^{\text{ième}}$ composante du vecteur α (correspondant à la $k^{\text{ième}}$ modalité MS) se calcule à plus basse résolution, entre la modalité $(B_{kl})^{interp}_0$ et la modalité A_l , version dégradée par convolution avec un filtre passe-bas de la modalité A_0 initiale. Donc :

$$\alpha_k(i, j) = (B_{kl})^{interp}_0(i, j) / A_l(i, j) \quad \text{Eq. 5.43}$$

Ensuite les auteurs forment le produit vectoriel entre les vecteurs B_0^* et B^{interp}_l , et démontrent que, d'après leurs hypothèses, ce produit vaut 0. La conclusion qui en découle est que ces vecteurs doivent être colinéaires (pour plus de détails concernant la démonstration, se référer à l'article de Garzelli *et al.* 2004). L'origine de cette méthode de fusion repose sur la distance *SAM* définie par ces mêmes auteurs (Alparone *et al.* 2003) et que nous avons décrit dans le chapitre 2. La fusion *SDM* consiste à minimiser l'angle *SAM* suivant :

$$SAM(B_0^*(i, j), (B_l)^{interp}_0(i, j)) = \arccos \left(\frac{\langle B_0^*(i, j), (B_{kl})^{interp}_0(i, j) \rangle}{\|B_0^*(i, j)\|_2 \| (B_{kl})^{interp}_0(i, j) \|_2} \right) \quad \text{Eq. 5.44}$$

Les auteurs concluent que la signature spectrale du produit fusionné est une fonction proche de celle de l'image suréchantillonnée. Une critique majeure peut être immédiatement faite : l'opération de suréchantillonnage peut se faire de manière très différente. Par exemple, un suréchantillonnage par réplication des pixels assure qu'aucune information spectrale erronée n'est introduite, en dépit d'une très mauvaise qualité géométrique. D'autres opérateurs offrent des compromis entre l'injection d'information tout en limitant les défauts géométriques : bilinéaire, bicubique, bicubique convolué par la *FTM* du capteur... Les images obtenues par ces opérateurs sont toutes différentes. Par conséquent, il existe autant de méthodes *SDM* différentes qu'il existe d'opérateurs de suréchantillonnage différents.

L'origine de cette imprécision est la formulation de Wang (2005) qui utilise la modalité originale interpolée comme point de départ de la fusion. Cette étape engendre une distorsion qui doit être soulignée et prise en compte dans le résultat. La méthode de fusion *SDM* ne répond pas à la propriété fondamentale des produits de fusion puisqu'elle vise à faire ressembler l'image fusionnée à l'image originale suréchantillonnée. Ainsi, le produit fusionné obtenu est proche de l'ensemble originale et ne reflète pas ce que le capteur correspondant aurait pu observer s'il possédait la résolution spatiale du Pan *res0*. Nous ne retenons pas cette méthode de fusion et n'avons pas souhaité tester ses performances lors des deux campagnes d'évaluation que nous avons organisées.

5.4.3 Le ratio 4 dans le cas particulier d'un MSM dyadique

Nous avons vu que dans ce cas, l'application du *MSM* fournissait des plans à des résolutions intermédiaires qui ne figurent pas dans les notations précédentes. Ceci concerne les couples *IMM-HRIMM M2*, *M3* et *RWM*, puisque les *IMM* des autres couples présentés (*M1*, *CBD* et *SDM*) ne nécessitent pas de se placer à ces résolutions plus basses que *res1*.

La démarche pour les trois couples *IMM-HRIMM M2*, *M3* et *RWM* est identique, à savoir que l'estimation des paramètres de l'*IMM* s'effectue entre les plans de détails situés entre *res1* et *2res1*. Le

HRIMM est ensuite appliqué à deux reprises, pour synthétiser les plans de détails situés entre $2res0$ et $res1$, et $res0$ et $2res0$.

5.5 Conclusion à propos des différents types de méthodes

5.5.1 Méthodes de type projection-substitution et de contribution spectrale relative

L'avantage de ces méthodes est de présenter une augmentation notoire de la qualité visuelle au niveau des structures présentes dans les images (Shettigara 1992, Zhou *et al.* 1998, Vijayaraj *et al.* 2004). La qualité géométrique des images synthétisées par les méthodes *IHS*, *Brovey*, *P+XS*, ... est jugée satisfaisante dans la plupart des publications. Ces méthodes sont ainsi adaptées à certaines applications, telles que des applications liées à la cartographie (Vijayaraj *et al.* 2004), ou encore pour la localisation d'un phénomène particulier ou d'une cible (Yocky 1996). Cependant, leur défaut majeur est la distorsion radiométrique qui a pour incidence une modification plus ou moins importantes des signatures spectrales, c'est-à-dire des couleurs des modalités MS synthétisées (Wald *et al.* 1997). Cette observation a été soulignée par Pellemans *et al.* (1993) qui avancent que ces méthodes ne sont pas adaptées à l'étude de la végétation.

La distorsion radiométrique est due à une modification des basses fréquences des modalités MS d'origine (Shi *et al.* 2005), et donc à un non-respect de la propriété de cohérence. Si on rappelle l'expression de la transformation *IHS* comme exemple de méthode de fusion appartenant à la première catégorie, et celle de Brovey pour la seconde :

$$B_{k0}^* = (B_{k1})^{interp}_0 + \delta, \text{ avec } \delta = Pan - I \quad (\text{Eq. 5.4})$$

$$B_{k0}^* = (B_{k1})^{interp}_0 \frac{A_0}{\sum_{k=1}^N (B_{k1})^{interp}_0} \quad (\text{Eq. 5.11})$$

On remarque que dans les deux situations, la modalité Pan est injectée dans son intégralité. Les basses fréquences de la modalité A_0 se retrouvent dans B_{k0}^* et viennent donc modifier celles de B_{k1} . De plus, dans le cas des méthodes de contribution spectrale relative, les basses fréquences d'une modalité B_k subissent aussi l'influence des autres modalités B_k par l'intermédiaire de la combinaison linéaire de ces modalités placées en dénominateur.

De plus, pour apporter de meilleurs résultats, ces méthodes nécessitent une forte corrélation entre chaque modalité B_k et la modalité A (Wald 2000). En effet, nous avons soulevé en introduction les différentes contraintes imposées par ce cadre particulier de la synthèse d'images B_k à haute résolution spatiale, à savoir les dissimilitudes entre A et B_k telles que les inversions de contraste et les occultations d'objets. Comme ces types de méthodes sont globales, elles ne sont pas en mesure de proposer une solution lorsque les contenus spatiaux des deux images considérées ne sont localement pas corrélés.

Le point faible des méthodes de la première catégorie vient du fait que toutes les modalités B sont traitées conjointement, avec une approche vectorielle. A partir du moment où une transformation d'espace intervient, la moindre modification apportée à la composante comportant la quantité maximale de structures, engendre un impact sur chacune des modalités B par l'intermédiaire de la transformée inverse. Cet impact diffère selon la projection d'espace sélectionnée (Zhou *et al.* 1998). Dans le cas de la méthode *IHS* linéaire, l'erreur produite est identique pour les trois modalités traitées. En ce qui concerne *ACP*, cet impact est fonction de la proportion dans laquelle chaque bande se retrouve dans le premier facteur. Certaines modalités peuvent concentrer l'impact principal du traitement alors que d'autres pourront ne subir que des modifications bénignes.

Le problème majeur des méthodes de type contribution spectrale relative est l'hypothèse de départ. La figure 5.8 reprend le graphique déjà montré en introduction à propos des réponses spectrales

normalisées des différentes modalités du satellite Quickbird. A ce graphique ont été ajoutés des rectangles grisés pour attirer l'attention sur les zones problématiques du spectre.

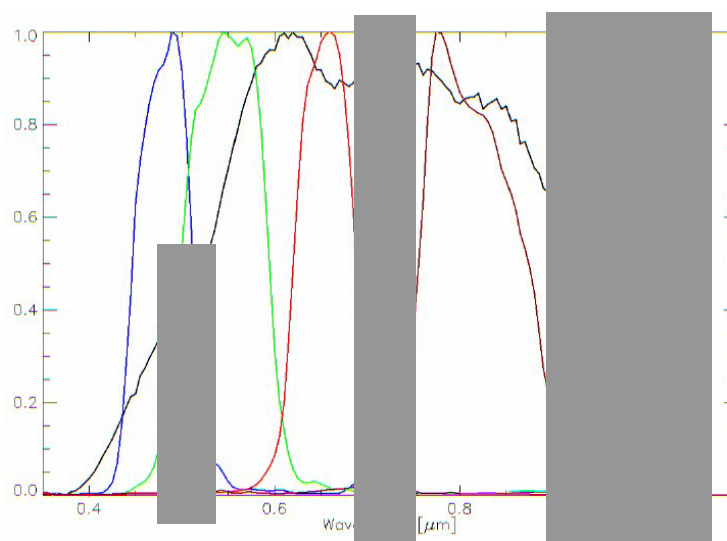


Figure 5.8 : réponses spectrales normalisées des canaux bleu, vert, rouge, PIR et Pan pour le satellite Quickbird, en fonction des longueurs d'ondes incidentes.

Les deux rectangles de droite rappellent que le capteur Pan est susceptible d'être sensible, par exemple, à un objet qui réfléchirait la lumière du soleil dans une longueur d'onde proche de $1\text{ }\mu\text{m}$, se manifestant par des niveaux de gris non nul dans cette modalité. Par contre, comme cette partie de spectre électromagnétique n'est couverte par aucune autre modalité MS, l'hypothèse de contribution spectrale relative n'est pas vérifiée. De plus, certaines modalités se chevauchent, créant dans ces deux images verte et bleue une redondance spectrale de l'information. Toutefois, cette relation est presque vraie dans le cas de la méthode *P+XS* développée pour les satellites SPOT1-4. Cette méthode a l'avantage de procéder dans le domaine des luminances, et utilise une spécification initiale du satellite, à savoir que le capteur Pan a été construit pour présenter des luminances qui soient égales à la moyenne des luminances des modalités XS1 et XS2. L'avantage de cette méthode n'est plus conservée lorsque l'on s'intéresse à un autre satellite. De plus, la modalité XS3 du satellite SPOT présente une bande passante totalement disjointe de celle de la modalité Pan, et la méthode *P+XS* suggère une simple duplication pour la modalité XS3. La duplication par la méthode des plus proches voisins a pour avantage de ne pas changer le contenu spectral d'une image. Cependant, même si cette technique respecte la propriété de cohérence, le résultat visuel est médiocre, et le suréchantillonnage ne peut pas être considéré comme un résultat de fusion à part entière.

5.5.2 Méthodes de type ARSIS

En ce qui concerne les méthodes de type *ARSIS*, la principale qualité est le respect de la propriété de cohérence, et donc une bonne restitution des signatures spectrales. En effet, les spectres des produits fusionnés, une fois ramenés à la résolution initiale, sont généralement fidèles à ceux des modalités *B* d'origine. Cependant, cette bonne restitution s'effectue au détriment des structures géométriques de l'image. Les images fusionnées présentent des artefacts tels que le halo, les rebonds, le bavé... qui sont décrits au chapitre suivant.

La qualité de synthèse de ce type de méthodes est fortement liée à la qualité de la réponse fréquentielle du filtre *MSM* utilisé pour dissocier les différentes fréquences spatiales du spectre. Si on reprend l'exemple de la méthode de *Pradines*, la modalité A_1 a été synthétisée en utilisant un filtre moyennneur correspondant à une fonction échelon, c'est-à-dire un sinus cardinal dans le domaine de Fourier. Cette fonction décroît progressivement en se rapprochant de la fréquence de coupure et produit de nombreux rebonds à l'origine d'un repliement spectral ou aliasing important. L'aliasing produit une modification des basses fréquences, et peut ainsi contribuer à créer une distorsion radiométrique du produit fusionné.

5.5.3 Rappel des précautions à prendre concernant la qualité d'une méthode

Nous voulons souligner ici les précautions à prendre lorsque l'on conclue quant à la qualité d'une méthode de fusion. Les conclusions doivent être nuancées puisque la qualité des produits fusionnés par une méthode dépendent :

- ◆ de l'écart entre les instants d'acquisition,
- ◆ des résolutions spatiales des images d'entrée,
- ◆ du ratio entre ces images,
- ◆ des réponses spectrales des capteurs (spectres disjoints, chevauchement, recouvrement total...),
- ◆ du contenu spatial et spectral de la scène,
- ◆ et enfin, de l'importance accordée par l'utilisateur aux différents aspects de la qualité des produits fusionnés.

5.6 Critique des méthodes de fusion récentes

Nous avons cherché dans la littérature récente les solutions trouvées par les auteurs afin d'atteindre à la fois une meilleure qualité spatiale et spectrale. Il existe plusieurs méthodes hybrides qui utilisent à la fois une approche multiéchelle combinée à une des méthodes des deux premières catégories. Parmi elles, nous avons accordé une attention particulière à la méthode de Ballester *et al.* (2003a, 2003b, 2006) car ses produits fusionnés présentaient une très bonne qualité de synthèse au niveau du bâti. Plusieurs autres méthodes de fusion récentes sont décrites. De plus, un regard critique est apporté quant aux dernières tendances de développement choisies vis-à-vis des contraintes imposées par la synthèse haute résolution des modalités MS.

5.6.1 La méthode de Ballester *et al.* (2003)

Il s'agit d'un modèle variationnel qui est le fruit d'une collaboration entre le centre national d'études spatiales (CNES) de Toulouse et l'université de Pompeu Fabra de Barcelone. Cette méthode ne fait partie d'aucune catégorie de manière évidente puisque les termes qui composent la fonctionnelle traitent différents aspects des produits fusionnés. Nous avons apporté un soin particulier à l'étude de cette méthode de fusion sur laquelle la DGA et le CNES avaient attiré notre attention car les travaux des auteurs montraient une bonne qualité géométrique des images présentées dans leurs communications et articles. De plus, elle a été définie en théorie pour respecter la propriété de cohérence et pour offrir une bonne restitution des hautes fréquences spatiales. En effet, un des trois termes est la retranscription mathématique de la propriété de cohérence monomodale. La bonne restitution géométrique provient d'un second terme qui s'appuie sur l'hypothèse que la géométrie des images B est contenue dans la carte topographique de l'image A . Les auteurs définissent la carte topographique par une image de gradients directionnels qui leur permettent de n'agir qu'au niveau des petites variations locales, ou hautes fréquences de l'image. Enfin, le troisième et dernier terme part de la même hypothèse que les méthodes de contribution spectrale relative puisque A est écrit comme une somme pondérée des modalités B .

Nous avons enrichi la plate-forme en IDL de cette nouvelle option de fusion. Nous avons eu l'opportunité d'établir des échanges scientifiques avec deux de ses auteurs, Vicent Caselles et Coloma Ballester de l'université de Pompeu Fabra. Ces diverses rencontres nous ont permis de connaître les diverses implantations de leur code et les limites, ainsi que d'envisager avec eux plusieurs améliorations.

La fonctionnelle d'énergie à minimiser est composée de trois termes : un terme lié à la géométrie basée sur le calcul de la carte topographique de A , un terme radiométrique et enfin, un terme d'attache aux données qui correspond à la propriété de cohérence monomodale. Ses minima donnent les images B^* reconstruites à haute résolution.

5.6.1.1 Le terme d'attache aux données

Un produit fusionné B_{k0}^* dégradé à sa résolution spatiale d'origine $(B_{k0}^*)_I$ doit être proche de sa modalité correspondante de départ B_{kI} . La modalité originale est l'image sans aucun pré-traitement, sauf spécification contraire. Le terme d'attache aux données $E_{attacheData}$ est un scalaire défini par :

$$E_{attacheData} = \sum_{n=1}^N \sum_{i,j=1}^{L1,L2} \left(\Pi_S \left(k_k * B_{k0}^*(i,j) \right) - B_{kI}(i,j) \right)^2 \quad \text{Eq. 5.45}$$

La modalité fusionnée B_{k0}^* est dégradée spatialement par convolution avec k_k pour la ramener à la résolution spatiale initiale $resI$. Notons que cette convolution s'effectue dans le domaine de Fourier pour augmenter la rapidité d'exécution. L'opérateur Π_S produit un sous-échantillonnage en sélectionnant un pixel tous les $resI/res0$ pixels dans chaque direction. A $resI$, on forme ensuite la différence pixel à pixel au carré avant de sommer les résultats sur tous les pixels qui décrivent la modalité originale de taille (L_1, L_2) , et sur toutes les modalités.

5.6.1.2 Le terme géométrique

L'hypothèse de départ est que la géométrie des modalités B_k est contenue dans la carte topographique de l'image A . La carte topographique est définie comme une fonction du gradient de cette image. Si on pose ϕ l'angle entre les gradients de A et de la modalité B_k et si « . » correspond au produit scalaire, alors :

$$\nabla A \cdot \nabla B_k = |\nabla A| |\nabla B_k| \cos(\phi) \quad \text{Eq. 5.46}$$

L'hypothèse se traduit par le fait que les gradients ∇A et ∇B_k doivent être colinéaires, c'est-à-dire que ϕ doit tendre vers 0 à π radians près. Alors :

$$\nabla A \cdot \nabla B_k = |\nabla A| |\nabla B_k| \quad \text{Eq. 5.47}$$

et :

$$\nabla A / |\nabla A| \cdot \nabla B_k = |\nabla B_k| \quad \text{Eq. 5.48}$$

On définit θ comme étant le gradient de A normé. Alors l'équation devient

$$\theta \cdot \nabla B_k = |\nabla B_k| \quad \text{Eq. 5.49}$$

Pour obtenir un terme à minimiser, on utilise l'orthogonal de θ noté θ^\perp . Ainsi, les gradients ∇A et ∇B_k doivent être orthogonaux.

$$\theta^\perp \cdot \nabla B_k = 0 \quad \text{Eq. 5.50}$$

Comme cette expression est identique à π radians près, elle est capable de traiter les inversions de contraste. Le terme géométrique proposé est :

$$E_{géométrique} = \sum_{k=1}^N \frac{\gamma_k}{N} \sum_{i,j=1}^{H1,H2} \left| \theta^\perp(i,j) \cdot \nabla B_{k0}^*(i,j) \right|^2 \quad \text{Eq. 5.51}$$

Une fois le produit scalaire formé entre les deux gradients, le résultat est élevé au carré pour chaque pixel avant d'être sommé sur les pixels de l'image fusionnée de taille (H_1, H_2) comme défini dans les notations. Un facteur de pondération γ_k permet d'assigner des poids différents pour chaque modalité avant de sommer le résultat sur le nombre total de modalités.

5.6.1.3 Le terme radiométrique

En terme de radiométrie, il est supposé que A est une combinaison linéaire des B_k , approche similaire aux méthodes de type « contribution spectrale relative ». D'où :

$$A_0 = \sum_{k=1}^N w_k B_{k0}^* \quad \text{Eq. 5.52}$$

où les w_k sont positifs ou nuls et leur somme est égale à 1. Le terme d'énergie correspondant est :

$$E_{radiométrique} = \sum_{i,j=1}^{H1,H2} \left(\sum_{k=1}^N w_k B_{k0}^*(i,j) - A_0(i,j) \right)^2 \quad \text{Eq. 5.53}$$

La différence pixel à pixel entre la somme pondérée des B_k et de A est calculée. Ensuite, de la même façon que pour les autres termes de la fonctionnelle, le résultat pour chaque pixel est élevé au carré avant de procéder à la somme sur tous les pixels et toutes les modalités B_k .

5.6.1.4 Fonctionnelle d'énergie et minimisation par la descente de gradient

L'énergie finale de la $k^{\text{ième}}$ modalité MS s'écrit :

$$E_k = \mu E_{\text{attacheData}} + E_{\text{géométrique}} + \lambda E_{\text{radiométrique}} \quad \text{Eq. 5.54}$$

où l'on va pouvoir jouer sur les valeurs de μ et λ pour contrôler l'influence des différents éléments de la fonctionnelle.

Les énergies géométriques et radiométriques se calculent à haute résolution spatiale $res0$. Par contre, l'énergie d'attache aux données est calculée à basse résolution. Cette fonctionnelle d'énergie est dérivée, et ses minima donnent les images fusionnées. La minimisation s'effectue par la méthode de la descente de gradient :

$$(B_{k0})^{p+1} = (B_{k0})^p + t(\nabla E_k)^p \quad \text{Eq. 5.55}$$

où p correspond à l'indice d'itérations de la minimisation, ∇E_k au gradient de l'énergie de la $k^{\text{ième}}$ modalité. Le pas t de la descente de gradient peut soit être fixé au départ, soit s'adapter au cours des itérations pour s'orienter dans la direction du gradient de E maximal.

Cette formulation impose que l'expression de ∇E_k donne des images qui possèdent la taille des images fusionnées (H_1, H_2). Le terme d'attache aux données est donc problématique puisqu'il œuvre à plus basse résolution spatiale, et la taille de l'image synthétisée par ce terme est plus petite (L_1, L_2) que celle attendue. On ne peut pas travailler à basse résolution car on ne tirerait pas profit des deux autres énergies de l'équation qui n'ont d'intérêt qu'à haute résolution. Ce terme a donc fait l'objet de plusieurs implantations différentes.

Dérivée du terme d'attache aux données :

J'ai noté la présence d'une erreur dans l'écriture de ce terme, qui me fut confirmée par Vicent Caselles. La version corrigée du terme dérivé fait apparaître les transposées du filtre k_k noté k_k^t et du peigne de Dirac Π_S^t , omis dans l'article de Ballester *et al.* (2003). La nouvelle expression est :

$$(\nabla E_{\text{attacheData}})_k = \Pi_S^t k_k^t * (\Pi_S(k_k * B_{k0}) - B_{kl}) \quad \text{Eq. 5.56}$$

Le terme entre parenthèses est identique à celui de la fonctionnelle. L'image résultant du calcul entre parenthèses est de la taille des images B originales, et donc plus petite que celle des images fusionnées. C'est le terme $\Pi_S^t k_k^t$ qui doit procéder au suréchantillonnage en créant une image de taille (H_1, H_2) comportant l'information liée au respect de la moyenne locale des images. Plusieurs manières d'implanter ces opérateurs ont été étudiées, et elles ont une forte influence sur le résultat ; quatre interprétations différentes de l'opérateur $\Pi_S^t k_k^t$, intitulés M_0, M_1, M_2 et M_3 , ont été implantées et testées par les auteurs.

La première implantation M_0 est la plus simple. L'image originale B_{kl} n'a subi aucun suréchantillonnage. Le terme résultant de la parenthèse est une image de taille égale à celle de l'image originale. M_0 consiste donc à voir l'opérateur $\Pi_S^t k_k^t$ comme un suréchantillonnage suivi d'une convolution. La petite image obtenue est agrandie en insérant le nombre de zéro nécessaire pour que sa taille corresponde à celle des images fusionnées. Ensuite, le filtre k_k^t est convolué au résultat. Ce filtre est identique au filtre k_k si le celui-ci est symétrique comme c'est le cas pour nos applications numériques.

L'implantation M_1 diffère du modèle précédent dans la manière de rééchantillonner la petite image intermédiaire. Au lieu d'insérer des zéros, les pixels de l'image sont répliqués avant la convolution par k_k^t .

Les implantations M_2 et M_3 appliquent la convolution avant l'étape du suréchantillonnage. Le filtre k_k^t est différent du filtre k_k , puisque celui-ci s'applique sur des images de taille égale à celle des images B originales. Les filtres utilisés pour la dégradation des modalités fusionnées sont de taille 41

par 41 pixels, et sont différents pour chaque modalité. Ces filtres sont calculés pour le type d'images usuellement traitées par les auteurs : Pélican et Quickbird. Le ratio entre les modalités initiales est donc de 4. Les auteurs créent des filtres de taille 5 par 5 pixels à partir des originaux. C'est l'opérateur Π_S^t qui diffère pour les deux implantations. Pour M_2 , une fois la convolution par ce filtre 5 par 5 appliquée, on réplique les pixels comme pour M_1 . Le suréchantillonnage pour M_3 s'effectue par l'utilisation d'un filtre supplémentaire qui pondère les valeurs du bloc 4 par 4 pixels en fonction de la position de chaque pixel dans ce bloc. Les résultats de ces diverses implantations sont présentées dans la partie suivante.

Dérivée du terme géométrique :

La dérivation théorique du terme géométrique dépend de l'approximation discrète adoptée pour le gradient. Ici :

$$\nabla^{+,+} f = (\nabla_x^+ f, \nabla_y^+ f) \quad \text{Eq. 5.57}$$

$$\nabla^{+,-} f = (\nabla_x^+ f, \nabla_y^- f) \quad \text{Eq. 5.58}$$

$$\nabla^{-,+} f = (\nabla_x^- f, \nabla_y^+ f) \quad \text{Eq. 5.59}$$

$$\nabla^{-,-} f = (\nabla_x^- f, \nabla_y^- f) \quad \text{Eq. 5.60}$$

avec :

$$\nabla_x^+ f(i, j) = f(i+1, j) - f(i, j) \quad \text{Eq. 5.61}$$

$$\nabla_y^+ f(i, j) = f(i, j+1) - f(i, j) \quad \text{Eq. 5.62}$$

$$\nabla_x^- f(i, j) = f(i, j) - f(i-1, j) \quad \text{Eq. 5.63}$$

$$\nabla_y^- f(i, j) = f(i, j) - f(i, j-1) \quad \text{Eq. 5.64}$$

Cette approximation discrète implique une réécriture de la composante géométrique de l'énergie :

$$E_{\text{géométrique}} = \sum_{k=1}^N \frac{\gamma_k}{N} \sum_{\alpha, \beta = +, -}^{H1, H2} \sum_{i, j=1} \left| \theta_{\alpha, \beta}^\perp(i, j) \cdot \nabla B_{k0}^{*, \alpha, \beta}(i, j) \right|^2 \quad \text{Eq. 5.65}$$

Pour la dérivation, on fait intervenir les opérateurs duaux de $\nabla^{+,+}$, $\nabla^{+,-}$, $\nabla^{-,+}$ et $\nabla^{-,-}$ qui sont respectivement les opérateurs de divergence $\text{div}^{-,-}$, $\text{div}^{-,+}$, $\text{div}^{+,-}$ et enfin $\text{div}^{+,+}$. La dérivée du terme géométrique est :

$$(\nabla E_{\text{géométrique}})_k = \frac{\gamma_k}{N} \sum_{\alpha, \beta = +, -} \text{div}^{\alpha*, \beta*} \left(\left\langle \theta_{\alpha, \beta}^\perp \cdot \nabla B_{k0}^{*, \alpha, \beta} \right\rangle \theta_{\alpha, \beta}^\perp \right) \quad \text{Eq. 5.66}$$

où α^* correspond au signe opposé de α et de même pour β . Pour faciliter l'implantation de cette expression, on explicite l'une des quatre divergences. Le produit scalaire entre $\langle \rangle$ se développe de la manière suivante :

$$\theta_{+,+}^\perp = 1/\sqrt{((\nabla_y^+ A)^2 + (\nabla_x^+ A)^2)} (\nabla_y^+ A, -\nabla_x^+ A) \quad \text{Eq. 5.67}$$

$$(\nabla B_{k0}^*)^{+,+} = (\nabla_x^+ B_{k0}^*, \nabla_y^+ B_{k0}^*) \quad \text{Eq. 5.68}$$

Pour simplifier l'écriture, soit (a, b) les coordonnées de $\theta_{+,+}^\perp$ et (c, d) celle du gradient $\nabla B_{k0}^{*,+,+}$. Alors, le produit scalaire est donné par :

$$\langle \theta_{+,+}^\perp \cdot (\nabla B_{k0}^*)^{+,+} \rangle = ac + bd \quad \text{Eq. 5.69}$$

qui est un coefficient multiplicateur que l'on appellera C pour chaque coordonnée de la divergence. Alors,

$$\text{div}^{-,-} (\langle \theta_{+,+}^\perp \cdot (\nabla B_{k0}^*)^{+,+} \rangle \theta_{+,+}^\perp) = C (a(i, j) - a(i-1, j) + b(i, j) - b(i, j-1)) \quad \text{Eq. 5.70}$$

Dérivée du terme radiométrique :

Il s'agit de la dérivation d'un terme linéaire :

$$(\nabla E_{\text{radiométrique}})_k = w_k \left(\sum_{j=1, j \neq k}^N w_j B_{j0}^* - A_0 \right) \quad \text{Eq. 5.71}$$

En conclusion, la bonne qualité visuelle des éléments du bâti provient de la corrélation élevée entre A et chaque modalité B_k sur ces zones. De plus, puisque l'un des termes de la fonctionnelle correspond à la transcription mathématique de la propriété de cohérence monomodale et que son poids

est élevé par rapport aux autres termes, les produits fusionnés offrent généralement une bonne cohérence avec les signatures spectrales des modalités originales.

Par contre, même si l'avantage du terme géométrique est d'être capable de traiter les inversions de contraste, il va avoir tendance à créer du bruit ou contraire du flou en présence de corrélation locale insuffisante. De plus, le terme radiométrique correspond exactement à l'hypothèse des méthodes de type contribution spectrale relative, par conséquent, les produits fusionnés vont aussi partiellement présenter les défauts de ce type de méthodes. Cette méthode de fusion a été testée et comparée à d'autres méthodes de fusion lors d'une campagne d'expérimentation. Les résultats de cette expérimentation sont placés dans le chapitre suivant.

5.6.2 Méthodes combinant IHS à ARSIS ou à la contribution spectrale relative

5.6.2.1 IHS et transformée en ondelettes

La lutte contre la distorsion spectrale de la méthode IHS a mené Nunez *et al.* (1999), Chibani et Houacine (2002) et Gonzalez-Audicana *et al.* (2004) à considérer seulement les hautes fréquences spatiales de A . Pour ne pas générer de modification des basses fréquences de la modalité B_k d'intérêt, ces auteurs utilisent la transformée en ondelettes « à trous » pour extraire les plans d'ondelettes de A . Si on se place, une fois encore, dans le cas par exemple du ratio 2 avec une transformée dyadique, qui permet de passer directement d'une résolution à l'autre par une application unique de l'algorithme ($res0-res1$, $res1-res2$...), la nouvelle intensité s'écrit alors comme la somme de l'intensité d'origine rééchantillonnée et du ou des plans d'ondelettes situés entre les résolutions spatiales de A et des modalités B d'origine :

$$I' = I + D_{A,0-1} \quad \text{Eq. 5.72}$$

L'avantage de cette méthode est de ne pas distordre l'ensemble original en produisant des couleurs fidèles à celles des compositions colorées originales. De plus, l'impression visuelle de l'image obtenue est très bonne, puisque l'intégralité des structures de A est injectée dans les produits fusionnés. Cependant, l'existence de dissimilitudes entre A et chaque modalité B_k n'est pas prise en compte. Cette approche n'utilise pas d'approche locale pour vérifier si une structure doit réellement se retrouver dans la nouvelle intensité. Si on reprend l'exemple de l'introduction du chemin qui est présent dans la modalité Pan à haute résolution et absent de la modalité bleue, cette méthode fera apparaître dans le produit fusionné bleu les structures hautes fréquences du chemin, c'est-à-dire les contours qui le délimitent.

5.6.2.2 IHS et prise en compte de la contribution spectrale relative des modalités

Tu *et al.* (2004) ont, par cette publication, ouvert une voie de développement reprise plusieurs fois par la suite. Le principe fondamental est de considérer l'écriture de l'intensité originale de départ de la transformée IHS linéaire :

$$I = 1/3 ((R_I)^{interp}_0 + (V_I)^{interp}_0 + (B_I)^{interp}_0) \quad \text{Eq. 5.73}$$

et de diminuer l'écart A et les modalités B en terme de contenu spectral, et de se rapprocher ainsi des méthodes de la catégorie contribution spectrale relative.

La première avancée a été de s'intéresser aux comportements des comptes numériques en présence de végétation.

En effet, dans le IHS original, la distorsion spectrale est encore plus flagrante sur la végétation qui possède une forte réponse dans l'infrarouge, couvert par le Pan mais non par la relation définie par l'équation 5.73 (figure 5.9). L'intérêt d'insérer la modalité PIR dans l'équation contribue à augmenter la corrélation entre le Pan et cette intensité en présence de végétation, et limite ainsi partiellement la distorsion spectrale. La nouvelle intensité définie par Tu et ses collaborateurs s'écrit comme la moyenne des quatre modalités MS où le poids de chaque modalité reste équivalent :

$$I = 1/4 ((PIR_I)^{interp}_0 + (R_I)^{interp}_0 + (V_I)^{interp}_0 + (B_I)^{interp}_0) \quad \text{Eq. 5.74}$$

Cette méthode a été appelée *Fast IHS (FIHS)* ou *Generalized IHS (GIHS)*. Notons que l'équation étant linéaire, elle peut aisément se généraliser à N bandes.

La seconde avancée concerne le chevauchement des modalités bleue et verte de la figure 5.9. Comme expliqué lors de la critique des méthodes de type contribution spectrale relative, ce chevauchement a pour effet de créer une redondance spectrale de l'information relative à une scène dans ces deux modalités. Pour prendre en compte le recouvrement partiel de ces modalités, les auteurs proposent de rectifier l'équation 5.71 en introduisant la pondération suivante :

$$I = 1/3 ((PIR_I)^{interp}_0 + (R_I)^{interp}_0 + a(V_I)^{interp}_0 + b(B_I)^{interp}_0) \text{ avec } a + b = 1 \text{ et } a < b \quad \text{Eq. 5.75}$$

Pour Ikonos, ils ont cherché sur un lot de quatre-vingt douze extraits d'images Ikonos à maximiser la corrélation entre le Pan et l'intensité I en faisant varier ces deux paramètres. La réponse était optimale pour le couple $(a, b) = (0.75, 0.25)$. Le reste de la procédure de fusion par *IHS* demeure inchangée ; A remplace alors cette nouvelle intensité avant de procéder à la transformation inverse pour revenir à l'espace initial de représentation des couleurs.

Dans la même optique de développement, Gonzalez-Audicana *et al.* (2006) ont conservé l'approche *GIHS* à quatre modalités, et ont proposé une nouvelle équation pour la nouvelle intensité :

$$I' = \frac{1}{4} \gamma A_0 \quad \text{Eq. 5.76}$$

où le coefficient γ est défini à partir de la proportion de spectre (surface) de chaque modalité B_k par rapport à la réponse spectrale de A .

La méthode de Choi (2006) correspond à une transformation partielle de l'intensité. En effet, I' s'écrit :

$$I' = A_0 - (A_0 - I)/t \quad \text{Eq. 5.77}$$

où le paramètre t permet de doser le compromis d'injection entre I et la modalité A .

Tous ces travaux ont cherché une issue au problème de la distorsion radiométrique de la transformée *IHS* en tentant de prendre en compte la contribution spectrale relative de chaque modalité avec A . Leur point commun est de s'accorder autour de l'expression proposé par Choi (2006) :

$$\text{Min } \{|A - I'|^2 + |I' - I|^2\} \quad \text{Eq. 5.78}$$

Or nous ne sommes pas d'accord avec cette formulation. D'une part, cette expression pousse la nouvelle intensité I' à ressembler à A , et à cause des dissimilitudes entre A et B déjà soulignées à plusieurs reprises dans ce chapitre, ces deux modalités peuvent localement être différentes. D'autre part, le second terme tend à diminuer l'écart entre l'intensité à basse résolution I , même si elle est rééchantillonnée, et celle à haute résolution I' . L'une comporte des hautes résolutions, l'autre non ; elles ne peuvent donc pas se ressembler.

Notons que la transformée *IHS* n'est pas la seule méthode de type projection-substitution à avoir été combinée avec des méthodes de type contribution spectrale relative. En effet, la méthode de fusion de Aiazzi *et al.* (2006) récemment développée, s'appuie sur une transformée de *Gram-Schmidt* de l'espace des couleurs et utilise, elle aussi, une relation de pondération des modalités B par rapport à A .

Il s'agissait pour eux du moyen pour exprimer le compromis entre qualité spectrale et qualité spatiale. Encore une fois, nous n'adhérons pas à cette idée de compromis. Même si les méthodes actuelles tendent à favoriser l'un ou l'autre de ces deux aspects, l'objectif de la synthèse d'images B à haute résolution spatiale ne doit pas aboutir à un choix d'une qualité au détriment d'une autre. Des précautions doivent être prises dans la formulation mathématique du problème pour ne pas restreindre les capacités de développement. Nous pensons qu'une méthode de fusion performante doit être en mesure d'atteindre une bonne qualité radiométrique et géométrique simultanément.

5.6.3 Retour à la définition de la fusion pour définir les voies de développement prometteuses

Nous nous sommes placés dans le cadre de la synthèse d'images MS à haute résolution spatiale qui soient les plus proches possibles de celles qui auraient été acquises par le capteur correspondant s'il possédait cette haute résolution. La synthèse à haute résolution s'effectue au moyen

d'une modalité possédant cette haute résolution spatiale, dans notre cas, il s'agit de la modalité Pan. L'information qui sépare la modalité MS d'origine et son homologue fusionnée correspond aux haute fréquences spatiales, qui sont présentes dans la modalité Pan.

Nous avons montré que la distorsion radiométrique est causée par une modification des basses fréquences de la modalité MS d'origine, comme l'avait aussi expliqué Shi *et al.* (2005). Les méthodes de type projection-substitution et contribution spectrale relative ont, par construction, tendance à distordre les basses fréquences. Cette distorsion est d'autant plus importante que la corrélation entre les modalités MS et le Pan est faible. Les méthodes ARSIS ne sont pas complètement protégées de ce défaut, et la distorsion sera dans ce cas dépendante de la réponse fréquentielle du filtre de décomposition utilisé.

La méthode de fusion idéale devrait vérifier la propriété de cohérence, c'est-à-dire ne pas modifier les basses fréquences de ce signal, et n'agir que sur les hautes fréquences spatiales. Si on se réfère à l'équation de la propriété de cohérence monomodale, la modalité fusionnée ramenée à sa résolution spatiale initiale $(B_{k0}^*)_I$ doit être égale à cette même modalité avant fusion à une petite erreur ε près. Alors :

$$(B_{k0}^*)_I = B_{kI} + \varepsilon \quad \text{Eq. 5.79}$$

Dans le domaine de Fourier, cette relation s'écrit :

$$TF(B_{k0}^*)_I = TF(B_{kI}) + TF(\varepsilon) \quad \text{Eq. 5.80}$$

La transformée de Fourier de l'erreur ε doit tendre vers zéro ce qui signifie que ε ne contient que des hautes fréquences ; pour vérifier la propriété de cohérence, l'algorithme de fusion doit préserver les basses fréquences du signal et apporter des hautes fréquences spatiales. Par conséquent, la fusion doit permettre de séparer le spectre en fréquences pour être capable de protéger les informations des échelles de résolutions égales ou inférieures à la résolution des images MS initiales. Pour cela, il est nécessaire d'avoir recours à des outils capables de procéder à une séparation efficace des fréquences ; par efficace, on entend des filtres qui ne produisent pas trop d'artefacts de type repliement spectral dans l'espace de Fourier.

Les outils permettant la séparation des fréquences sont les approches multiéchelles et multirésolution. Ils appartiennent à un cadre mathématique rigoureux permettant de se déplacer et d'agir à des échelles différentes. Les contraintes sur les filtres dépendent de l'algorithme utilisé. En ce qui concerne la littérature, il a été trouvé que Yocky (1996), Wald *et al.* (1997), Zhou *et al.* (1998), Wald (2002), Ranchin *et al.* (2003), Wang *et al.* (2005) et Shi *et al.* (2005) soutiennent que l'approche multiéchelle permet d'établir le compromis entre le respect des basses fréquences et l'insertion des hautes fréquences. Mangolini (1994) et Yocky (1996) ont montré que les méthodes de fusion utilisant la transformée en ondelettes permettaient de préserver les caractéristiques spectrales des images en limitant la distorsion de l'ensemble MS original.

En conclusion, nous avons procédé à une étude critique de la littérature par rapport aux contraintes imposées par le cadre de la fusion d'images que l'on s'est fixé. Les faiblesses et les points forts de chaque méthode ont été soulevés. L'étude des voies de développement récentes a mené à la définition du cadre de développement de nos futures méthodes de fusion. Notre choix s'est porté sur les méthodes de type ARSIS qui utilisent des filtres dont la réponse en fréquence satisfait un certain nombre de contraintes. Nous n'avons cependant pas restreint notre champ d'investigation à ce type de méthodes.

VI Nouvelles méthodes de fusion

Il a été rapporté à plusieurs reprises que les méthodes de fusion de type *ARSIS* donnaient de bons résultats statistiques et une bonne cohérence avec l'ensemble MS d'origine (Couloignier *et al.* 1998, Ranchin et Wald 2000, De Boissezon et Laporterie 2003, Ranchin *et al.* 2003, Alonso-Reyes *et al.* 2005a). La bonne cohérence avec l'ensemble original implique un respect des signatures spectrales contenues dans les modalités MS originales lors de l'augmentation de la résolution spatiale. La raison principale du respect de la propriété de cohérence par les méthodes de type *ARSIS* est que les noyaux de convolution utilisés sont réguliers au premier ordre, c'est-à-dire que la moyenne du plan de détails introduit est égale à 0.

L'analyse quantitative des produits de fusion renseigne sur de grandes distorsions, comme une injection insuffisante d'énergie dans les hautes résolutions spatiales avec une différence en variance entre le produit fusionné et sa référence négative, ou une distorsion radiométrique générale avec un biais non nul. Par rapport à l'approche qualitative, elle a l'avantage d'être sensible par exemple à une variation d'un compte numérique sur l'ensemble des pixels, que même un œil averti ne peut pas détecter.

L'évaluation quantitative de la qualité ne peut toutefois pas déceler tous les artefacts des produits de fusion. La complexité de la perception humaine nous procure des atouts essentiels dans la distinction des défauts lorsqu'une image est synthétisée par un processus de traitement d'image quelconque. Il n'existe actuellement pas de distance capable de reproduire cette complexité pour pouvoir nous affranchir de l'analyse qualitative. Ces deux approches sont totalement complémentaires. En dépit d'une qualité statistique évidente et d'une bonne restitution générale des couleurs, les implantations du concept *ARSIS* sont multiples, et la qualité visuelle diffère fortement d'une réalisation à l'autre. L'injection des structures de hautes fréquences peut être trop forte ou trop faible, entraînant des artefacts visuels tels que le flou, du bruit, ou encore des contours trop rehaussés. Les couleurs peuvent aussi être localement mal synthétisées ou baver sur les pixels alentours (De Boissezon et Laporterie 2003, Ranchin *et al.* 2003). Ces artefacts sont néfastes à l'interprétation et aux post-traitements (classification, détection d'objets...).

Dans le but de développer de nouvelles méthodes de fusion, nous avons choisi la stratégie d'identifier visuellement des artefacts des méthodes de type *ARSIS*. Ces nouvelles connaissances guideront les développements de nouvelles méthodes qui tenteront de préserver la qualité quantitative, tout en limitant les artefacts. L'identification et l'explication constituent la première partie de ce chapitre. Cinq implantations *ARSIS*, parmi les plus performantes à ce jour, ont été appliquées sur dix images en vraies et fausses couleurs provenant de trois satellites différents (SPOT5, Ikonos et Quickbird) et de simulations CNES. Les produits de fusion ont été confiés à quatre interprètes d'images lors d'une première campagne d'analyse visuelle dont le mode opératoire a été placé en annexe (annexe 2). Les résultats de cette analyse conjointement à nos observations ont permis la sélection d'une série d'extraits pour illustrer les différents types d'artefacts recensés. On apporte ensuite une explication mathématique sur l'origine de ces défauts.

Les conclusions de cette étude ont mené au développement de trois nouvelles méthodes de fusion. Elles ne font pas seulement appel aux méthodes de type *ARSIS*, et puisent dans d'autres catégories de méthodes des solutions pour éviter de reproduire ou au moins atténuer les artefacts. La qualité visuelle et statistique de ces nouvelles méthodes de fusion a été établie au moyen de l'application du protocole de validation proposé au chapitre 2 de ce document de thèse. Pour l'analyse

visuelle à haute résolution spatiale des produits de fusion, nous avons pu bénéficier d'une nouvelle intervention des interprètes d'images. Plusieurs méthodes existantes sont mises en concours avec les nouvelles méthodes de fusion pour montrer la performance des nouvelles méthodes.

6.1 Artefacts des méthodes de type ARSIS

On présente ici les cinq types d'artefacts que nous avons pu observer visuellement sur des résultats issus de différentes méthodes de fusion de type ARSIS. Cette partie est complétée par l'explication de l'origine des divers défauts observés.

6.1.1 Méthodes de fusion testées et illustrations choisies

Les méthodes de fusion sélectionnées ont été testées grâce à l'outil logiciel fusion sous IDL (annexe 1). Leurs éléments ont déjà été présentés.

- ◆ *ATWT-M3* qui utilise comme *MSM* la transformée à trous combinée au couple *IMM-HRIMM M3* qui s'appuie sur une estimation globale des paramètres entre plans de coefficients d'ondelettes.
- ◆ La seconde méthode de fusion testée est *ATWT-RWM*. Au *MSM* « à trous » est associé le modèle local *RWM* pour l'estimation des paramètres. La taille de la fenêtre d'estimation a été fixée à 17 par 17 pixels.
- ◆ *GLP-CBD* utilise la décomposition hiérarchique de la pyramide Laplacienne généralisée *GLP* conjointement avec le modèle d'estimation locale *CBD* des paramètres qui s'applique aux approximations. La taille de la fenêtre d'estimation est fixée à 7 par 7 pixels.
- ◆ *GLP-M3* se réfère au même modèle multi-échelle que la méthode précédente et y joint le couple *IMM-HRIMM M3*.
- ◆ Enfin, *MallatDaub4-RWM* s'appuie sur le cadre de l'analyse multirésolution de *Mallat* avec l'ondelette *Daub4* directionnelle et non-symétrique, et utilise le couple *IMM-HRIMM RWM*. La décomposition avec cette ondelette produit une dégradation conséquente de la qualité visuelle des images. Nous l'avons choisi pour servir de comparatif dans cette première campagne. Cependant, comme la qualité a été jugée vraiment médiocre par les interprètes d'images, elle n'apparaîtra pas dans les commentaires et les illustrations.

Les images sur lesquelles ont été appliquées les algorithmes de fusion sont très diverses, en terme de satellites, résolutions spatiales, ratio et encore de paysages, comme en témoignent les cinq groupes d'illustrations suivantes des figures 6.1 à 6.5. Chaque figure est associée à une série de commentaires directement issus du rapport des interprètes d'images. Nous ferons référence à ces figures lors de la description des artefacts. Des cercles de couleurs différentes permettent de focaliser l'attention sur les artefacts. Nous avons recensé cinq grands types de défauts auxquels nous avons associé une couleur de cercle différente :

- ◆ le faïençage, avec des cercles de couleur rouge,
- ◆ les artefacts ponctuels ou irisations de couleurs aberrantes, avec des cercles verts,
- ◆ les inversions de contraste encadrés en bleu,
- ◆ les stries diagonales de radiométries successivement claires et sombres indépendantes du faïençage, défauts encadrés de blanc,
- ◆ enfin, le flou, entouré en rose.



a)



b)



c)



d)



e)

Figure 6.1 : image CNES, Toulouse - a) produit fusionné par *ATWT-M3*, b) par *ATWT-RWM*, c) par *GLP-CBD*, d) par *GLP-M3*, e) Pan. Copyright CNES 2000.

Sur ce premier extrait de l'image CNES Toulouse en vraies couleurs, les interprètes ont observé :

- ♦ par la méthode *ATWT-M3* : un léger flou sur l'ensemble de l'image,
- ♦ *ATWT-RWM* : des zones aléatoires de flou. Flou très important sur la végétation. On observe un phénomène de rebond sur les linéaires de la piste d'athlétisme (cercle blanc),

- ♦ *GLP-CBD* : des zones aléatoires de flou, qui est important dans les zones de transitions naturelles/artificielles. Accentuation des linéaires accompagnée d'un piqué et de quelques artefacts une fois encore sur la piste d'athlétisme et sur les pelouses,
- ♦ *GLP-M3* : un léger flou dans les zones homogènes (pelouse, terrain).



a)



b)



c)



d)



e)

Figure 6.2 : image Ikonos, Hasselt (Belgique) - a) produit fusionné par *ATWT-M3*, b) par *ATWT-RWM*, c) par *GLP-CBD*, d) par *GLP-M3*, e) MS originale.
Copyright Space Imaging 2002.

Observations des interprètes concernant la seconde illustration :

- ♦ *ATWT-M3* : des couleurs ternes,
- ♦ *ATWT-RWM* : une bonne restitution des couleurs, mais de nombreux artefacts, une haute radiométrie saturée et un rebond. Cet exemple met en évidence la dégradation importante réalisée par cet algorithme de fusion.
- ♦ *GLP-CBD* : des couleurs plus intenses, mais une modification de la texture du canal avec l'apparition d'une bande centrale de radiométrie claire.
- ♦ *GLP-M3* : une atténuation de la teinte des couleurs avec une perte quasi totale du vert. Apparition d'une bande centrale claire dans le cours d'eau.



a)



b)



c)



d)



Figure 6.3 : image Ikonos, centre ville de Fredericton (Canada) - a) produit fusionné par *ATWT-M3*, b) par *ATWT-RWM*, c) par *GLP-CBD*, d) par *GLP-M3*, e) MS originale. Copyright Space Imaging 2002.

e)

Observations des interprètes concernant la troisième illustration :

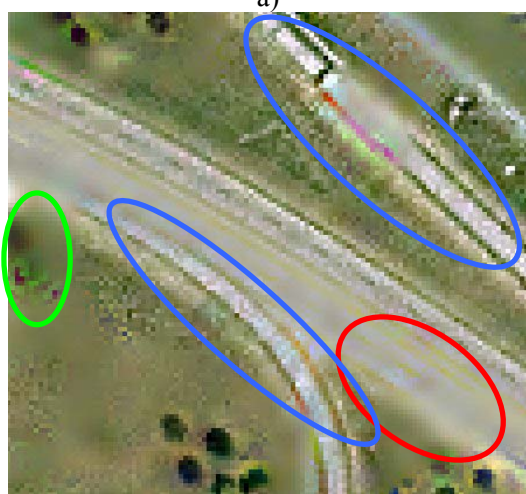
- ◆ *ATWT-M3* : flou général sur l'image,
- ◆ *ATWT-RWM* : phénomène d'irisation, c'est-à-dire artefacts ponctuels colorés (cercles blancs).
- ◆ *GLP-CBD* : phénomène d'irisation. Cet exemple montre des zones hétérogènes flagrantes (cercles rouges).
- ◆ *GLP-M3* : stries en diagonal sur le toit des bâtiments.



a)



b)



c)



d)



Figure 6.4 : image Ikonos, abords de Fredericton (Canada) - a) produit fusionné par *ATWT-M3*, b) par *ATWT-RWM*, c) par *GLP-CBD*, d) par *GLP-M3*, e) MS originale. Copyright Space Imaging 2002.

e)

Observations des interprètes concernant la quatrième illustration :

- ♦ *ATWT-M3* : bonne qualité visuelle, même si un peu lisse,
- ♦ *ATWT-RWM* : inversion de radiométrie des voies de circulation (ovale orange). Le bord des routes est crénelé, comme s'il avait été dessiné en pointillés. Un flou important occupe le milieu de la route,
- ♦ *GLP-CBD* : des tâches floues sont présentes sur la végétation et sur les routes (cercles blancs). Des artefacts de couleur longent la route ; présence de piqué de couleur verte dans la végétation. On observe la même inversion de contraste sur la route,
- ♦ *GLP-M3* : bonne impression visuelle de l'image, moins floue que celle obtenue avec *ATWT-M3*.



a)



b)



c)



d)



e)

Figure 6.5 : image Quickbird, centre ville de Fredericton (Canada) - a) produit fusionné par *ATWT-M3*, b) par *ATWT-RWM*, c) par *GLP-CBD*, d) par *GLP-M3*, e) MS originale. Copyright Digital Globe 2002.

Observations des interprètes concernant la cinquième illustration :

- ♦ *ATWT-M3* : les transitions entre les hautes radiométries des bâtiments et leurs ombres sont trop molles, perte de la couleur de l'objet jaune pâle en bas à droite de l'image,
- ♦ *ATWT-RWM* : l'image apparaît aussi un peu floue. Par contre, la couleur de l'objet en bas à droite se rapproche de celle attendue,
- ♦ *GLP-CBD* : le défaut principal est le crénelage, c'est-à-dire la succession de radiométries claires et sombres le long du toit du bâtiment,
- ♦ *GLP-M3* : même remarque que pour la méthode précédente.

6.1.2 Bilan de la qualité visuelle, méthode par méthode

Les interprètes d'images s'accordent à dire que la fusion a un apport informatif non négligeable pour toutes les résolutions et capteurs en leur possession lors de cette première expérimentation. La fusion permet de mieux discriminer les éléments du réseau routier, de la ville ou encore d'apprécier la nature d'un sol. Elle permet aussi de mieux caractériser les surfaces qui possèdent des radiométries voisines sur l'image Pan. Ils ont néanmoins conclu que les défauts sont moins flagrants pour les images ayant pour origine un capteur à résolution spatiale moyenne (SPOT5). Les conclusions des interprètes sont en accord avec les différents travaux publiés sur l'intérêt de la fusion.

Méthode *ATWT-M3* (image a des figures 6.1 à 6.5) : les images issues de cette méthode de fusion comportent un léger flou homogène sur toute l'image. Par exemple, la figure 6.4d présente une meilleure impression visuelle avec un contraste plus important que pour l'image 6.4a synthétisée par la méthode *ATWT-M3*. Des cercles de couleur rose permettent aussi de visualiser ce défaut. Par contre,

on observe une bonne similitude des couleurs comparées aux MS d'origine, mais elles apparaissent souvent un peu ternes. Les hautes radiométries sont souvent entourées d'un halo clair. La perception des petits objets est bonne, même si la couleur des petits objets a tendance à diffuser (baver) sur les pixels voisins. On note une légère atténuation des objets dans les ombres.

Méthode *ATWT-RWM* (image b des figures 6.1 à 6.5) : un flou important est présent sur toutes les images mais de manière aléatoire. Les zones de transition entre le bâti et la végétation, et plus généralement entre zones artificielles et naturelles, sont les plus dégradées. La couleur verte a tendance à baver. Excepté ce point, la restitution des couleurs est très bonne, particulièrement sur des objets de couleurs vives, tels que des voitures rouges ou des toits de bâtiments oranges (figure 6.3b), même si parfois la couleur contamine partiellement la radiométrie des pixels alentours. En conclusion, on note une bonne fiabilité des couleurs. Cependant, on a pu observer avec l'illustration 6.1b mais surtout 6.4b des inversions de contraste. Plusieurs irisations ponctuelles multicolores ont aussi été rapportées. Les hautes radiométries présentent des halos/rebonds. Des petits objets disparaissent par agrégation ou par le lissage des structures. On note une perte d'information dans les ombres.

Méthode *GLP-CBD* (image c des figures 6.1 à 6.5) : les remarques sont assez similaires à celles de la méthode *ATWT-RWM*. Néanmoins, les produits fusionnés ne présentent pas d'inversion de contraste. L'avantage aussi est que les couleurs apparaissent un peu plus intenses. Cependant, on note une présence accrue des artefacts ponctuels colorés et la présence de stries obliques dans des zones homogènes non floues.

Méthode *GLP-M3* (image d des figures 6.1 à 6.5) : les remarques ressemblent à celles obtenues pour la méthode de fusion *ATWT-M3*. Cependant, cette méthode de fusion présente moins de flou, mais plus de risques de faire apparaître des stries obliques (figure 6.5d).

En conclusion, les méthodes de fusion *ATWT-M3* et *GLP-M3* se révèlent juste acceptables pour un cadre opérationnel. Si la qualité informative de ces deux méthodes est acceptable, leur qualité visuelle n'est pas satisfaisante en raison d'un effet de voile et la présence de halos induits principalement par les hautes radiométries. La méthode *ATWT-M3* est considérée comme celle présentant la meilleure qualité visuelle.

6.1.3 Description et explications des artefacts visuels

6.1.3.1 Le faïençage (cercles rouges)

Le faïençage est le terme technique qui décrit les hétérogénéités locales. Il se manifeste par des zones floues distribuées de manière aléatoire dans l'image. Cet effet est visible dans tous les encadrés rouges des illustrations, et est exclusivement présent dans les illustrations des images fusionnées par les méthodes de fusion *ATWT-RWM* et *GLP-CBD*. L'intensité de ce phénomène s'accroît dans les zones de transitions entre la végétation et le bâti, c'est-à-dire de manière plus générale, entre les zones naturelles et artificielles.

L'origine du faïençage provient du choix des couples *IMM-HRIMM* pour ces deux méthodes de fusion, qui sont *RWM* et *CBD*. Ces couples ont pour point commun de procéder à une estimation locale des paramètres a et b , et à leur application conditionnelle à haute résolution en fonction de la valeur du coefficient de corrélation. Ce sont les objets visés par ces deux processus qui sont différents. Pour *RWM*, l'approche concerne les images de détails aux différentes résolutions, alors que *CBD* estime ses paramètres à partir des images d'approximation (*IMM*) et les applique au(x) plan(s) de détails (*HRIMM*). Si la corrélation spatiale locale est insuffisante entre Pan et MS (plans de détails ou contextes selon la méthode), alors les coefficients de détails à haute résolution spatiale sont fixés à 0. Des détails fixés à 0 ont pour effet de produire un suréchantillonnage des modalités MS originales, c'est-à-dire qu'aucune haute fréquence n'est injectée, alors que les pixels qui les entourent peuvent avoir des valeurs différentes de 0, et donc rehausser les structures. Un flou qui jouxte une zone fusionnée corrompt la régularité des structures d'une image, ce qui est très inconfortable pour la visualisation et pour les post-traitements comme une classification.

Par contre, même si l'origine du faïençage est le couple *IMM-HRIMM*, on remarque sur ces exemples que l'impression visuelle n'est pas parfaitement identique selon le *MSM ATWT* ou *GLP*. Dans le cas de l'implantation *ATWT-RWM*, ces zones donnent l'impression d'une peinture à la gouache sur laquelle on aurait fait tomber une goutte d'eau qui vient mélanger les tons. Pour cet algorithme de décomposition hiérarchique, toutes les images possèdent les mêmes dimensions. Ainsi, si des coefficients d'ondelettes sont égaux à zéro, la basse résolution spatiale est conservée, ce qui revient à recopier quelques pixels d'une version plus grossière et donc plus floue dans l'image fusionnée.

Par contre, pour la méthode *GLP-CBD*, les zones de flou contiennent des stries obliques, comme le montre la figure 6.3c (cercle rouge de droite). La figure 6.3e montre que les stries de la figure 6.3c sont semblables à celles présentes dans l'image MS d'origine suréchantillonnée. Les stries dans les zones de faïençage proviennent de l'utilisation de l'algorithme *GLP* qui est un *MSM* décimé. L'image MS d'origine présente une taille plus petite que l'image fusionnée. Une étape supplémentaire de suréchantillonnage est nécessaire pour adapter la taille de l'image à celle de l'image fusionnée finale, ce qui produit les stries obliques de la figure 6.3c, puisqu'il s'agit localement de la même opération que celle appliquée pour créer la figure 6.3e.

6.1.3.2 Les artefacts colorés (cercles verts)

Ce type de défaut est mis en valeur par des cercles de couleur verte sur les illustrations. La figure 6.2b montre dans le cercle de gauche un amas de pixels de couleur rose sur la piste d'athlétisme, et sur la droite, des pixels côte à côte qui présentent une radiométrie aberrante. De même, la figure 6.4c montre un arbre dont la couleur tend vers le vert fluo et les bâtiments de la figure 6.3b sont irisés, c'est-à-dire tachetés de pixels de couleurs vives.

L'origine de ce défaut est aussi inhérent aux méthodes locales *ATWT-RWM* et *GLP-CBD*, et indirectement au faïençage. La couleur d'un pixel de la composition colorée de l'image fusionnée dépend à la fois de la localisation de ce pixel dans l'image puisque ces méthodes suivent des approches locales, mais aussi de sa valeur dans chaque modalité MS. En effet, si la corrélation locale est insuffisante dans l'une des modalités MS et non dans les autres, alors la couleur de ce pixel aura tendance à tendre vers la couleur des modalités qui auront été fusionnées, au détriment de celle qui aura été suréchantillonnée. En bref, ces deux méthodes de fusion ne garantissent pas une homogénéité du résultat d'un point de vue multispectral. Cette hétérogénéité peut conduire localement à une mauvaise estimation locale qui peut se manifester par des valeurs de pixels avec une dominante de l'une ou l'autre modalité MS.

6.1.3.3 L'inversion de contraste (cercles bleus)

On remarque une première inversion de contraste sur la partie supérieure de la piste d'athlétisme de la figure 6.1a, alors que la partie inférieure ne présente pas cette inversion. Le résultat final apparaît irrégulier et très bruité. Les exemples 6.4b et c sont les manifestations les plus flagrantes de l'inversion de contraste. Les méthodes locales *ATWT-RWM* et *GLP-CBD* n'ont pas été en mesure de synthétiser correctement la route qui devrait apparaître plus sombre que la signalisation au sol, comme sur la modalité Pan.

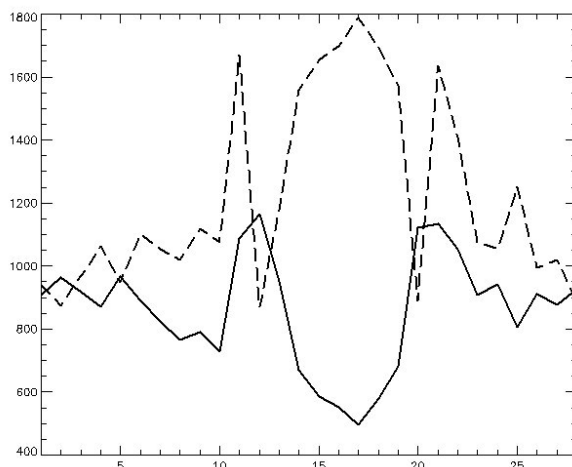
Nous avons cherché à comprendre le comportement local de ces processus de fusion. Les illustrations sont des compositions colorées produites à partir des modalités dont la bande spectrale appartient au visible. L'inversion de contraste n'apparaît pas dans les modalités PIR. Les figures 6.6a et b rappellent les deux compositions colorées de la figure 6.4. Les deux flèches rouges matérialisent les transects de ces deux images.



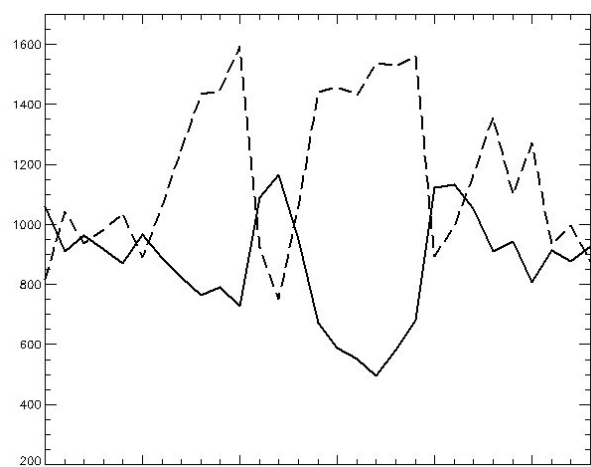
a)



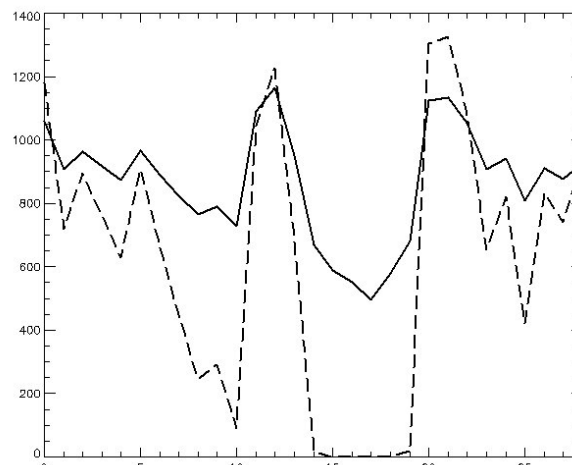
b)



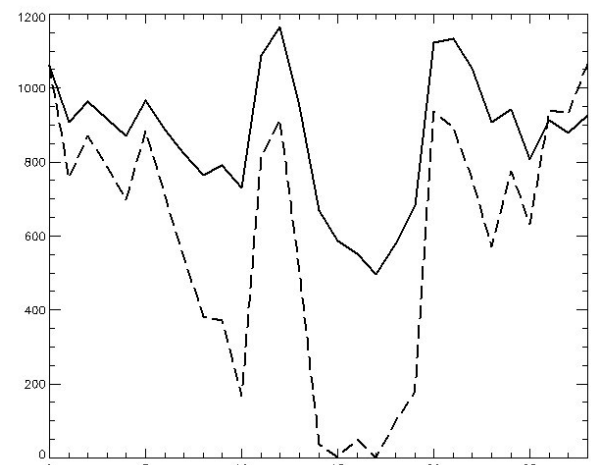
c)



d)



e)



f)

Figure 6.6 : image Ikonos, échangeur autoroutier, Fredericton (Canada). a) image couleur fusionnée par *ATWT-RWM*, b) image couleur fusionnée par *GLP-CBD*, c) courbe en trait plein, transect de l'image Pan correspondant à l'extrait de la ligne matérialisée par une flèche rouge dans les figures a et b, en pointillés, transect de la modalité rouge obtenue par *ATWT-RWM*, d) idem pour la modalité PIR fusionnée par *ATWT-RWM*, e) idem pour la modalité rouge fusionnée avec *GLP-CBD*, f) idem pour la modalité PIR, fusionnée avec *GLP-CBD*.

La modalité Pan sert de référence aux transects des images fusionnées. La courbe correspondante est en trait plein sur toutes les illustrations. Les figures 6.6c et d montrent en pointillés respectivement les modalités rouges fusionnées par les méthodes de fusion *ATWT-RWM* et *GLP-CBD*.

Le comportement de la modalité rouge est similaire à celui des autres modalités du visible, à savoir que la végétation et la route sont en inversion de contraste dans les deux images. Dans les deux cas, on remarque très nettement l'inversion de contraste au niveau de la route.

Par contre, comme nous l'avons dit auparavant, les deux méthodes de fusion locales ont été capables de synthétiser correctement la route pour les modalités PIR. Les transects Pan/PIR ont été placés figures 6.6e et f. En effet, les transects des modalités PIR sont en phase avec celui du Pan. La question dorénavant est de savoir pour quelles raisons les méthodes de fusion ont échoué dans la synthèse haute résolution des modalités visibles.

Intéressons-nous tout d'abord au comportement de la méthode *ATWT-RWM*. L'estimation des paramètres a et b s'effectue au niveau des plans d'ondelettes Pan et MS, respectivement figure 6.7a et b, situés entre les résolutions 4 et 8 m.

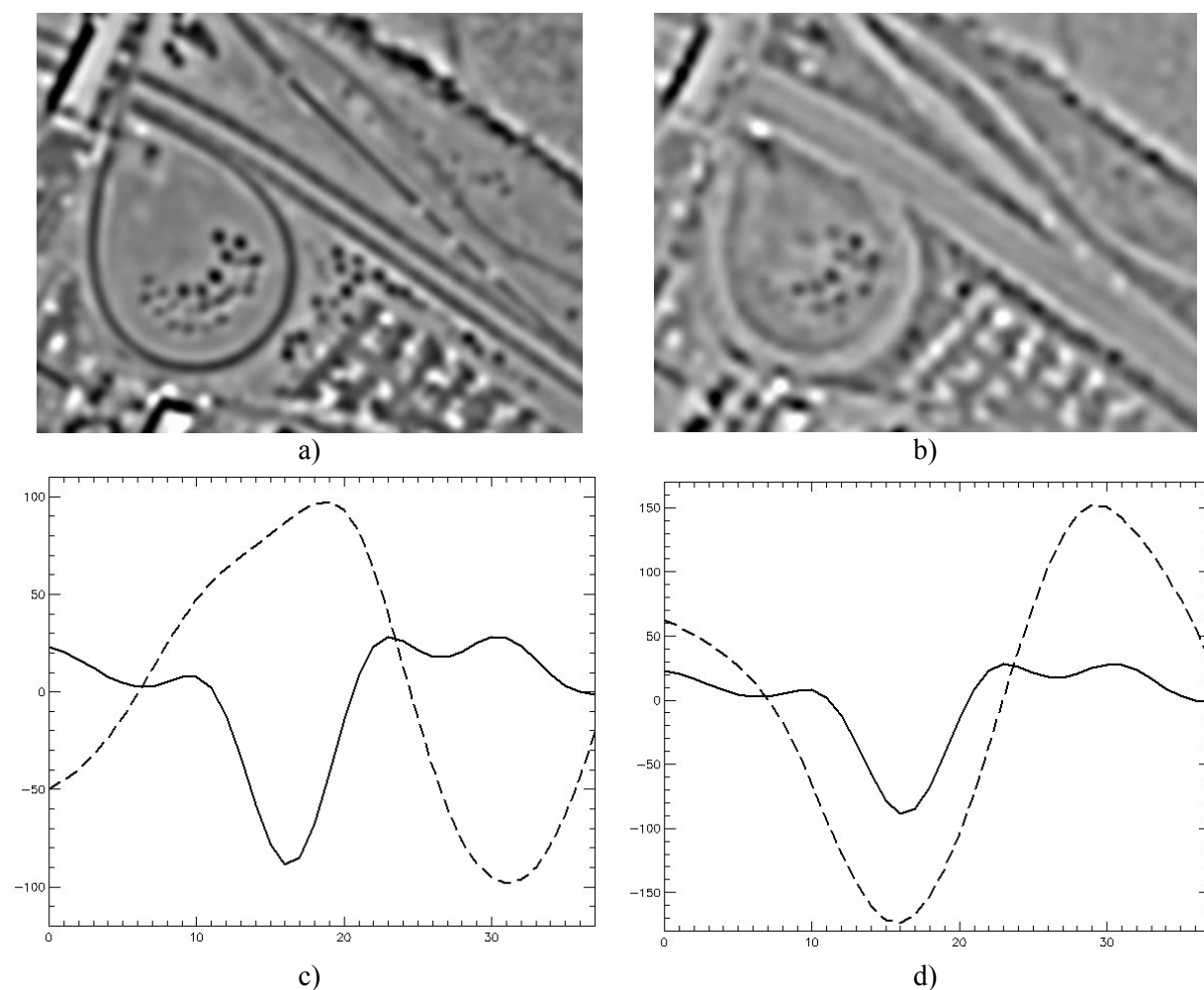


Figure 6.7 : image Ikonos, échangeur autoroutier, Fredericton (Canada). a) plan des coefficients d'ondelettes situé entre les résolutions spatiales 4 et 8 m pour la modalité Pan, b) idem pour la modalité rouge, c) transect Pan/modalité rouge, d) idem pour la modalité PIR.

La figure 6.7c montre les transects correspondant à ces deux images en trait plein pour l'image Pan et en pointillés pour la modalité rouge. On voit très clairement que ces deux courbes sont en inversion de phase avec une corrélation locale comprise entre -1 et -0.7 , alors que dans le cas de la modalité PIR placée en d, les courbes des détails entre le Pan et le PIR sont en phase au niveau du phénomène observé. La méthode globale *ATWT-M3* a correctement synthétisé la radiométrie de la route et de la signalisation pour les modalités dont la bande spectrale est située dans le visible. La valeur du coefficient global a pour la modalité rouge est de 0.95. Dans le cas du modèle local *ATWT-RWM*, les coefficients a pour la synthèse ont été estimés à des valeurs entre -1 et -2 au cœur de la

route, avec des valeurs de a d'un pixel à l'autre pouvant passer de -2 à 2.8 . Ce modèle n'assure à aucun moment une quelconque homogénéité spatiale des résultats sur le paramètre a . Les valeurs de a obtenues pour la modalité PIR valent 0.91 pour le modèle global, et entre 2.5 et 3 pour le modèle local, ce qui confirme que les détails ont été injectés avec une proportion différente mais avec un coefficient multiplicateur positif dans les deux cas. Les modalités Pan et PIR sont en phase car la bande passante du Pan contient une part d'infrarouge. La végétation qui a une forte réponse dans l'infrarouge apparaît plus claire que le bâti dans ces deux modalités.

Nous avons vérifié si l'on retrouvait ce comportement pour la méthode *GLP-CBD*. Dans cette méthode, l'estimation s'effectue sur les plans de coefficients d'approximation Pan et MS à l'échelle des modalités MS originales, c'est-à-dire à 4 m. La modalité Pan a été dégradée par l'algorithme *GLP* et placée figure 6.8a. La modalité rouge originale est placée figure 6.8b. La figure 6.8c donne les transects Pan/modalité rouge correspondant à ces deux extraits, et la figure 6.8d, les transects Pan/PIR.

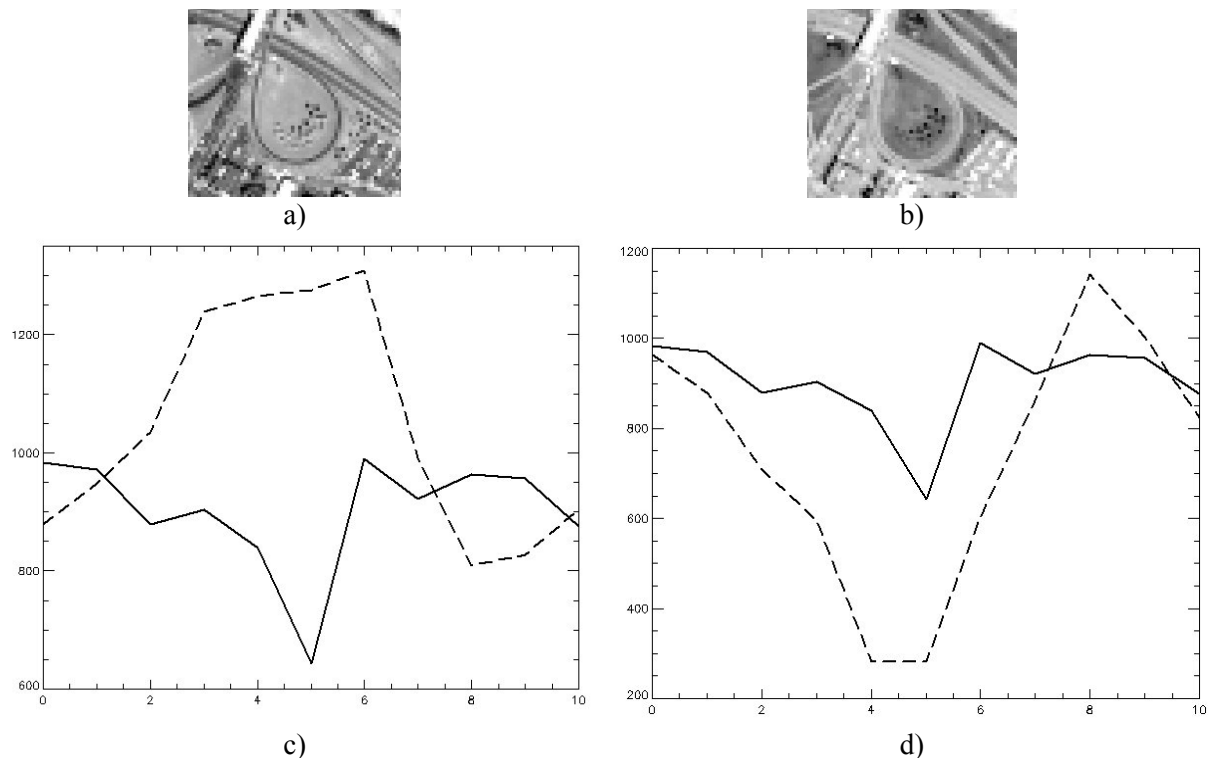


Figure 6.8 : image Ikonos, échangeur autoroutier, Fredericton (Canada). a) plan d'approximation Pan à 4 m, b) modalité rouge originale, c) transect Pan/modalité rouge, d) idem pour la modalité PIR.

Les courbes sont visiblement en antiphase dans le cas de la modalité rouge avec une anticorrélation forte puisque inférieure à -0.8 , et en phase dans le cas de la modalité PIR. On retrouve exactement le même comportement que pour la figure précédente. La méthode globale *GLP-M3* donne une valeur pour le paramètre a égale à 1.12 pour la modalité rouge, et 1.08 pour la modalité PIR. Ces valeurs proches de 1 signifient que les coefficients *GLP* haute résolution sont injectés presque sans aucune modification de leurs valeurs. Dans le cas de la synthèse locale par *CBD*, les coefficients le long de la route sont négatifs, comme l'on pouvait s'y attendre, avec une valeur comprise entre -1.5 et -2 . Ceux pour la modalité PIR atteignent la valeur du seuil supérieur (qui vaut 2) au lieu de 1.08 .

Dans les cas des deux méthodes de fusion locales, l'estimation s'effectue à une échelle où la signalisation de la route (lignes blanches) est noyée dans la radiométrie des pixels voisins ; l'échelle caractéristique de la signalisation est inférieure à la résolution spatiale d'estimation. Pour les tailles habituelles de fenêtre, de l'ordre de 11 par 11 pour le modèle *GLP* et 23 par 23 pour *ATWT*, l'anticorrélation est élevée pour les modalités visibles. Les méthodes de fusion ne sont pas en mesure de synthétiser la route et sa signalisation avec comme seule condition la corrélation.

Nous avons étudié l'influence de la taille de la fenêtre d'estimation sur la qualité de synthèse. Nous avons augmenté progressivement la taille de la fenêtre ; pour une fenêtre avec une taille de l'ordre de 29×29 pour le *MSM GLP*, la route est totalement floue, ce qui signifie que la corrélation a diminué en valeur absolue, et que sa valeur est inférieure au seuil en deçà duquel les modalités MS sont suréchantillonnées. Ce seuil vaut 0.37 pour la modalité rouge qui nous sert d'illustration. La corrélation locale est contrainte à la fois par l'inversion de contraste de l'environnement et de la route, et un comportement en phase des pixels situés à l'extérieur de cette zone problématique. En effet, pour que la route soit fusionnée avec une corrélation positive, il faut qu'une proportion suffisante de pixels en phase avec le Pan soit contenue dans la fenêtre d'estimation. Plus la taille de la fenêtre augmente, plus la zone située entre l'échangeur et les bâtiments de la ville, qui est correctement synthétisée, s'étend. La proportion de route fusionnée qui respecte la radiométrie du bitume et de la signalisation augmente. Les trois quarts de la route sont fusionnés pour une fenêtre de 79×79 , et son intégralité avec une fenêtre 99×99 . Comme on pouvait s'y attendre, lorsque la taille de la fenêtre tend vers la taille de l'image, la valeur des paramètres a et b estimés sur les fenêtres tend vers les valeurs des paramètres a et b globaux. En ce qui concerne l'algorithme *ATWT*, comme il n'est pas décimé, une taille 4 fois supérieure est nécessaire pour une synthèse correcte.

En apparence, dans les exemples cités ci-dessus, les méthodes de fusion locales présentent des artefacts qui affaiblissent les post-traitements appliqués aux produits fusionnés, et gênent l'interprétation. Le faïençage, les irisations et ces inversions de contraste ont contribué à ce que les interprètes rejettent systématiquement ces méthodes de fusion. On peut alors légitimement se demander quel est l'intérêt des méthodes de fusion exploitant la corrélation locale. La figure 6.9 suivante reprend l'exemple de l'inversion de contraste placé dans le premier chapitre. Il permet de montrer que dans certains cas, les méthodes locales peuvent offrir la possibilité de capter les inversions de contraste, et de manière plus générale, les dissimilitudes locales entre le Pan et les modalités à fusionner, contrairement aux méthodes globales.

6.1.3.4 Autre situation d'inversion de contraste

On rappelle que les images placées figure 6.9a et b correspondent à un extrait de la ville de Fredericton au Canada, acquis par le satellite Quickbird dans la modalité respectivement Pan et bleue suréchantillonnée par un opérateur bicubique. Les courbes placées en c correspondent aux sections des lignes sélectionnées par les rectangles rouges sur les deux images. L'inversion de contraste est évidente.

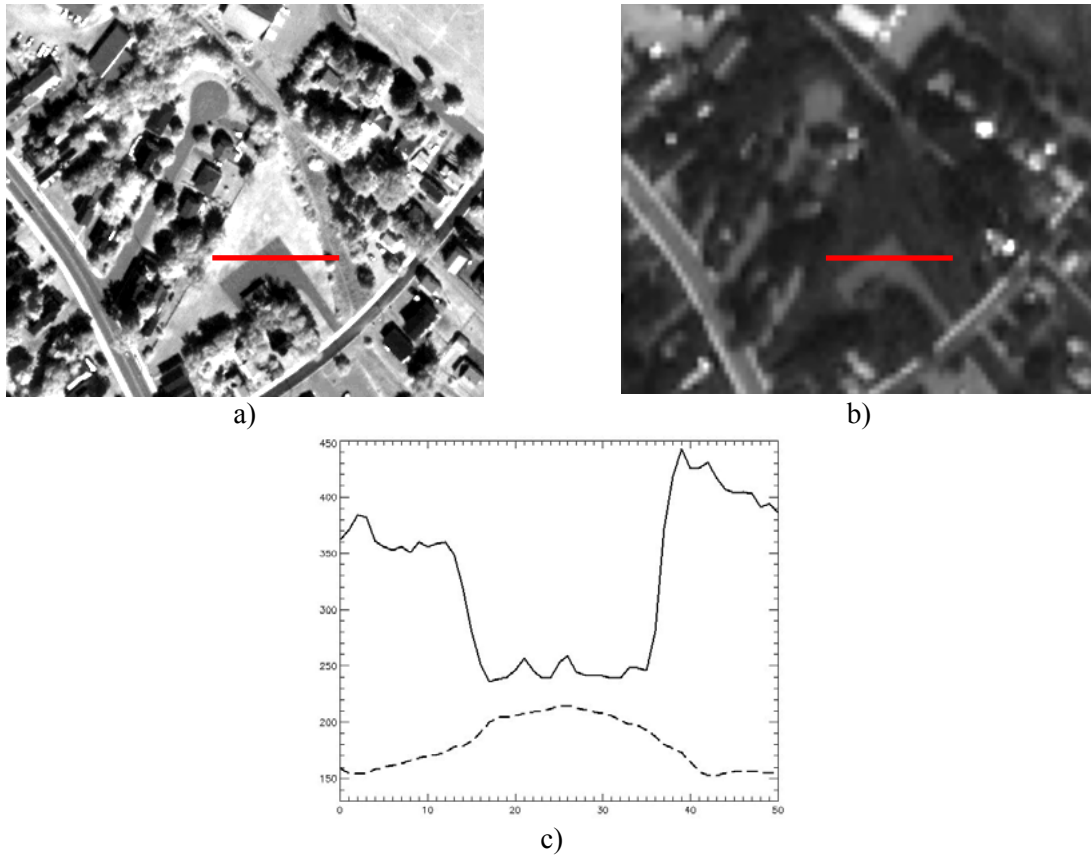


Figure 6.9 : image Quickbird, extraits de la zone périurbaine de Fredericton (Canada). a) Pan, 0.7 m, b) modalité bleue suréchantillonnée par un opérateur bicubique, 2.8 m, c) transects de ces deux images correspondant aux encadrés rouges. Copyright Digital Globe 2002.

La figure 6.10a et b représentent les produits de fusion par *ATWT-M3* et *ATWT-RWM*, et les graphiques c et d les transects correspondant à la même section que dans les images précédentes. Le trait plein représente encore une fois la modalité Pan, qui sert de comparatif.

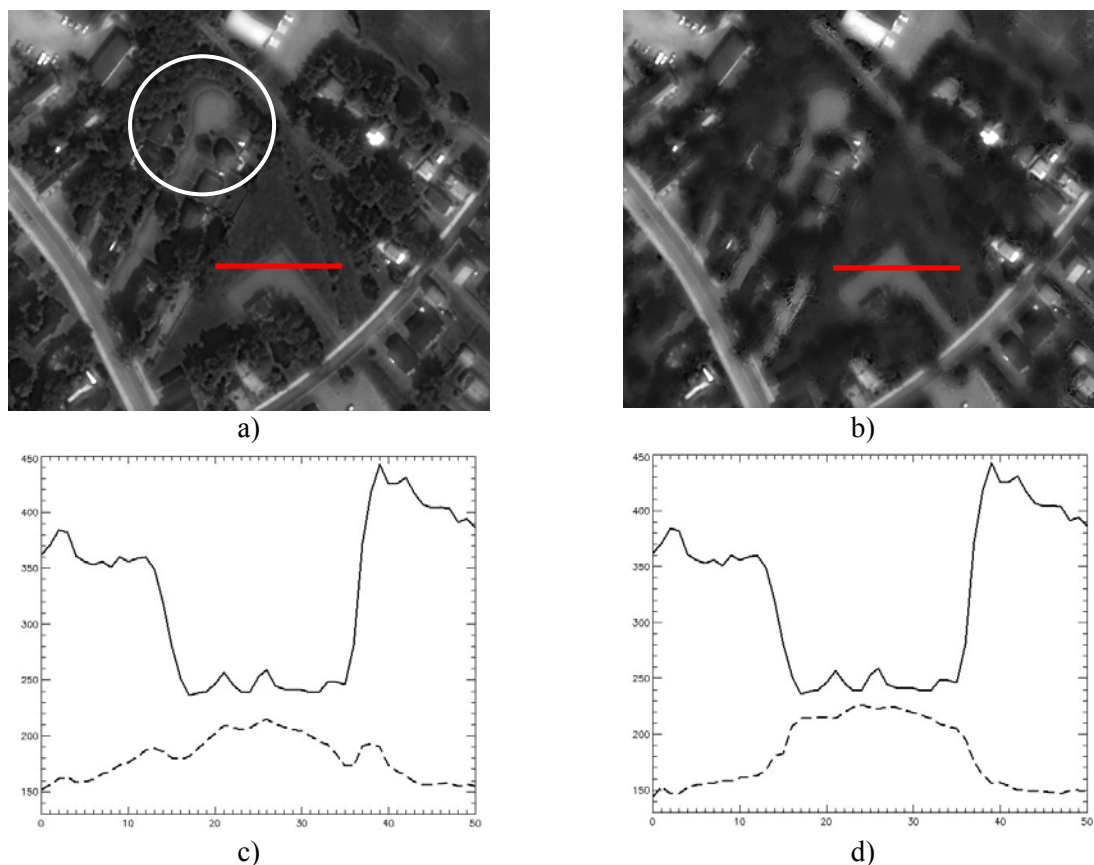


Figure 6.10 : a) modalité bleue fusionnée par *ATWT-M3*, b) fusionnée par *ATWT-RWM*, c) et d) en trait plein, le transect de la modalité Pan, et en pointillés, les transects respectifs des deux produits fusionnés.

La méthode de fusion globale *ATWT-M3* a généré un rebond erroné le long des contours, comme si le bâtiment était entouré d'un muret. Ce rebond est aussi présent un peu plus haut sur l'image au niveau du rond point (cercle blanc). Si on revient à la figure 6.9, on se rend compte que l'on obtient cet artefact dès lors que l'on se trouve sur des objets qui présentent des inversions de contraste avec leur environnement comparé à l'image Pan. Par contre, l'approche locale conditionnelle *RWM* a respecté le gradient de l'image bleue puisque les contours ont été rehaussés « dans le bon sens ». Une fusion qui repose sur une approche globale n'a pas été en mesure de traiter cette dissimilitude locale.

La différence avec l'exemple précédent est que l'inversion de contraste présente dans cet exemple est aussi présente à basse résolution spatiale. La taille caractéristique des objets en inversion de contraste (le rond point, le bâtiment...) est supérieure à la résolution spatiale où l'on procède à l'estimation des paramètres a et b , c'est-à-dire entre 2.8 et 5.6 m. En effet, les objets sont encore visibles à de telles résolutions spatiales.

6.1.3.5 Conclusion sur le défaut de type inversion de contraste

Le succès de la fusion vis-à-vis de ce défaut dépend de deux éléments : la taille caractéristique de l'objet par rapport à l'échelle d'estimation, mais aussi des propriétés spectrales de cet objet. Nous nous plaçons dans le cas où une plus grande proportion des pixels de la modalité MS sont en phase (c'est-à-dire corrélation positive) avec le Pan. Les conclusions sont susceptibles d'être différentes dans le cas contraire. Dans ce cas, la bande d'observation de cet objet qui le place ou non en inversion de contraste avec son environnement est prépondérante.

Dans le cas où un objet est de taille suffisamment importante, c'est-à-dire qu'il occupe encore quelques pixels à la résolution spatiale d'estimation des paramètres, deux situations sont possibles : si l'objet est en inversion de contraste et que la taille de la fenêtre utilisée dans l'approche locale est adaptée à la taille de l'objet, alors la méthode de fusion locale aura une chance de correctement

synthétiser cet objet. Par contre, l'approche globale va échouer. Il s'agit du cas de la partie précédente 6.1.3.4.

Dans le cas où la taille de l'objet est insuffisante, c'est-à-dire que l'objet a disparu à l'échelle d'estimation, il ne peut en aucun cas avoir une influence sur le calcul des paramètres. Par conséquent, l'approche globale prend en compte le comportement global de l'image que nous avons admis avec une corrélation positive entre le Pan et les modalités MS. Ainsi, si l'objet est en inversion de contraste, l'approche globale échouera, mais sera couronnée de succès dans le cas contraire. Les mêmes remarques peuvent être établies pour l'approche locale, qui prend en compte le comportement du voisinage ; si l'objet est en inversion de contraste comme le voisinage, alors la fusion par l'approche locale donnera un bon résultat, mais sera un échec dans le cas contraire. Ainsi, l'exemple de la partie 6.1.3.3 correspond au cas où l'objet (route + signalisation) est en phase avec l'image, le modèle global est un succès. Par contre les pixels avoisinants correspondent à de la végétation et sont localement en inversion de contraste dans les modalités visibles par rapport à la modalité Pan. L'approche locale échoue et introduit des coefficients de détails avec le mauvais gradient.

6.1.3.6 Les stries obliques hors faïençage

Ces stries sont différentes de celles rapportées lors de l'explication du faïençage dans le cas de la méthode de fusion *GLP-CBD*. Ce défaut est entouré d'un cercle blanc. Ainsi, la figure 6.3d, mais surtout les figures 6.5c et d montrent sur une zone relativement homogène représentant des toits de bâtiments, une succession de bandes de radiométrie plus ou moins sombres dans la direction diagonale. Le point commun des deux méthodes des figures 6.5c et d qui montrent ces stries est le *MSM GLP*.

L'illustration suivante permet de comprendre ce phénomène. La figure 6.11a montre une transition sélectionnée dans une image Pan acquise par le satellite Quickbird. Cette transition correspond au bord d'un bâtiment de la ville de Fredericton (Canada). Cet extrait est dégradé deux fois par l'algorithme non-décimé *ATWT*. Pour reconstruire à l'identique l'image Pan, il suffit d'ajouter les deux plans d'ondelettes correspondant à l'analyse. Au lieu de cela, nous avons multiplié l'écart-type de ces deux plans par 2 avant de les réinjecter. L'image obtenue a été placée en b. Elle apparaît beaucoup plus contrastée que l'image de référence. Elle donne un piqué erroné à l'image, avec un halo sombre autour des deux objets ponctuels en haut de l'image ainsi qu'un rebond de radiométrie opposée le long des structures : un rebond sombre lorsque la radiométrie de l'objet est claire, et inversement. Nous avons procédé de la sorte aussi pour obtenir l'image c, mais cette fois-ci, l'algorithme de décomposition hiérarchique de l'information utilisé est *GLP*.

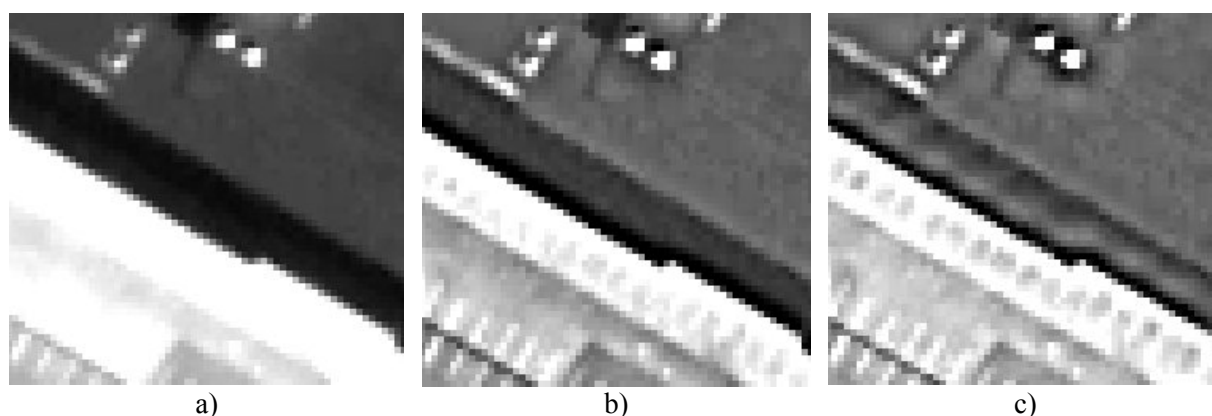


Figure 6.11 : image Quickbird, bord d'un bâtiment de Fredericton (Canada). a) extrait de référence provenant d'une image Pan, b) même image dégradée par *ATWT*, puis reconstruite avec ses deux plans de coefficients d'ondelettes de plus haute résolution spatiale multipliés par un facteur 2, c) idem mais avec l'algorithme *GLP*.

Copyright Digital Globe 2002.

La figure 6.11c montre une augmentation générale du contraste de l'extrait, des objets ponctuels de radiométrie élevée entourés d'un halo sombre. On remarque aussi dans l'ombre du bâtiment oblique sur l'image la présence des stries, qui sont similaires à celles des images des figures

6.5c et d. On a artificiellement créé ce défaut en injectant des coefficients de détails *GLP* dont la dynamique est trop élevée par rapport à celle nécessaire à la bonne reconstruction de la référence. On remarque que dans le cas d'un algorithme décimé, lorsque l'on fait une erreur en injectant des coefficients de détails dont la valeur est trop élevée, cette erreur se répercute à travers le suréchantillonnage par le biais de stries.

6.1.3.7 Le flou

Le flou a essentiellement été relevé dans les produits de fusion de la méthode *ATWT-M3* ; ce défaut est encadré en rose dans les illustrations 6.1 à 6.5. Les analyses de la propriété de synthèse sur différents produits de fusion résultant de cette méthode globale montrent que généralement, la variance du produit de fusion est inférieure à la variance de la référence. Cela signifie que l'on n'a pas injecté suffisamment d'énergie au niveau des hautes fréquences de l'image.

Nous avons procédé à une expérience similaire à la précédente, à savoir que nous avons modifié une fois encore la dynamique/variance des plans d'ondelettes lors de la reconstruction de l'image de référence. La figure 6.12 suit la même logique que la figure précédente, excepté que les coefficients ont été divisés par deux.

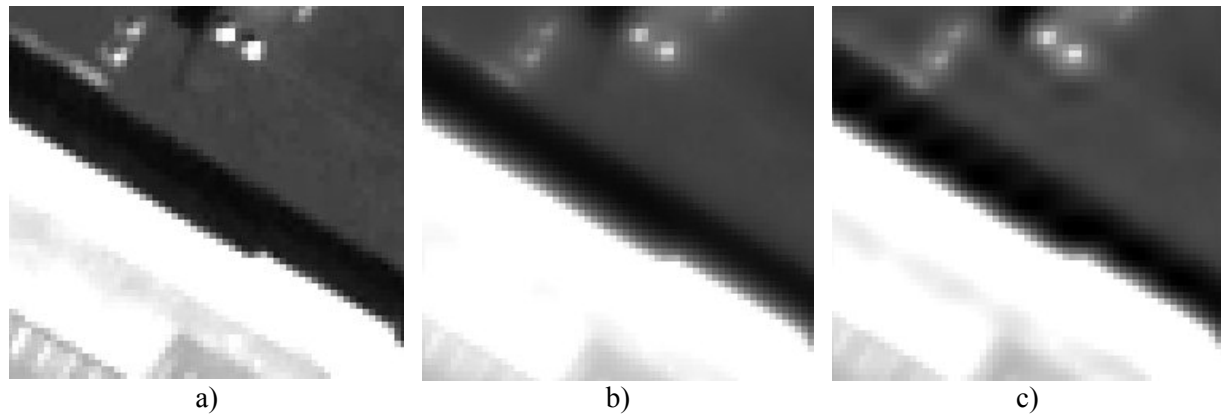


Figure 6.12 : image Quickbird, bord d'un bâtiment de Fredericton (Canada). a) extrait de référence provenant d'une image Pan, b) même image dégradée par *ATWT*, puis reconstruite avec ses deux plans de coefficients d'ondelettes de plus haute résolution spatiale divisés par un facteur 2, c) idem mais avec l'algorithme *GLP*.
Copyright Digital Globe 2002.

L'effet produit par ce traitement est un flou général, quelque soit le *MSM* utilisé. Les transitions sont plus progressives, les ponctuels en haut de l'image ont tendance à baver sur les pixels voisins.

Ceci montre que le flou est lié à un manque d'énergie injectée. Ce flou provient du fait que le modèle d'estimation des paramètres est un modèle global (méthodes *ATWT-M3* et *GLP-M3*). Comme nous l'avons expliqué précédemment, un modèle global ne peut prendre en compte les dissimilarités locales. Ces dissimilarités ont pour effet d'augmenter l'écart entre la modalité MS à fusionner et la modalité Pan, c'est-à-dire de diminuer la corrélation entre les images. Pour le couple *IMM-HRIMM M3*, l'estimation des paramètres s'effectue entre coefficients d'ondelettes Pan et MS. Que ce soit par une minimisation moindres carrés ou par centre d'inertie, le modèle *M3* cherche à tracer une droite par régression linéaire sur le nuage de points entre le Pan et chaque modalité MS. Mais comme la corrélation n'atteint pas sa valeur idéale de 1, les paramètres *a* et *b* estimés seront entachés d'erreur à cause des zones de corrélation moindre. Ils rentrent en contradiction avec le comportement local qui aurait donné une autre valeur aux paramètres. Si par contre la corrélation entre ces images est élevée, le modèle global (ici *M3*) donnera de meilleurs résultats.

6.1.4 Conclusion : vers de nouvelles méthodes de fusion

Les interprètes ont confirmé la présence d'artefacts géométriques qui avaient déjà été soulignés par le passé (De Boissezon et Laporterie 2003, Ranchin *et al.* 2003). Ces observations ont été aussi confirmées par les résultats du concours « fusion d'images » organisé par IEEE (Alparone *et al.* 2006). Notre outil d'évaluation a en effet été choisi par le comité d'organisation et notre laboratoire a procédé à l'évaluation des différents produits de fusion créés. Nous avons ainsi pu vérifier la présence de ces défauts pour d'autres images et d'autres méthodes de fusion. Quelques éléments de réponse ont été apportés quant à la formation de ces défauts lors des différents processus de fusion. La conclusion de ces analyses est que les deux approches, globale et locale, comportent des avantages et des inconvénients.

D'une part, les modèles globaux s'appliquent à synthétiser les structures géométriques en respectant leur régularité, au détriment d'un aspect un peu brumeux de l'image finale. Toutefois, le modèle global a l'avantage de procéder de manière équivalente sur toute l'image. Ainsi, même si un artefact est créé par la méthode, il est également distribué dans l'image, ou tout au moins peut s'observer sur tous les objets de même type. Ainsi, le flou de la méthode *ATWT-M3* et les stries obliques de la méthode *GLP-M3* sont présents sur toute l'image, et se visualisent plus facilement le long des transitions brutales de radiométrie.

D'autre part, les méthodes locales ont systématiquement été rejetées car l'aspect hétérogène du résultat se révèle trop inconfortable pour l'utilisateur et entraîne un manque de fiabilité en cas de post-traitements. En effet, contrairement aux approches globales, une zone de l'image peut subir un traitement totalement différent de ses pixels avoisinants, et ce traitement peut différer en fonction de la taille de la fenêtre d'estimation. Par contre, l'avantage des modèles locaux d'estimation est la qualité des BQ de leurs produits de fusion (cf tableaux 6.2 et 6.3 de ce chapitre). Cette qualité s'explique par le fait que l'injection de données du Pan vers les modalités MS ne s'effectue qu'en cas de corrélation 'suffisante' par rapport à un seuil donné. Dans le cas contraire, le modèle contourne la difficulté en procédant à un suréchantillonnage de la modalité originale basse résolution, donnant ainsi un meilleur respect de la propriété de cohérence avec l'ensemble original. En dépit de l'aspect hétérogène engendré par le suréchantillonnage, l'avantage est de s'intéresser au comportement local des images. Comme nous l'avons souligné en introduction, certains objets peuvent être occultés, voire même présenter une radiométrie contraire d'une image à l'autre. Les modèles globaux ne sont toujours pas capables de capter les dissimilitudes locales.

Voilà pourquoi nous avons développé avec Gabriel Peyré, chercheur au CNRS, une méthode de fusion exploratoire qui nous a permis d'analyser le comportement des approches locales conditionnelles exclusivement basées sur la corrélation locale en images de détails Pan et MS. La méthode de fusion est décrite avec précision en annexe (annexe 5). Le principe consiste à maximiser la corrélation locale en faisant varier différents paramètres tels que la pondération des pixels de la fenêtre d'estimation en fonction de leur distance au pixel central et la taille et l'orientation de cette fenêtre. Nous avons vérifié le bon comportement dans l'orientation des fenêtres d'estimation sur une photographie de visage placée aussi en annexe ; l'algorithme était bien en mesure de capter les directions de régularité maximale dans l'image. Par contre, l'évaluation visuelle des résultats à haute résolution spatiale a montré des hétérogénéités se manifestant avec des zones de flou plus ou moins denses. En suivant une arête de bâtiment, une partie pouvait tout à fait avoir été correctement rehaussée alors qu'une autre tendait à disparaître. La conclusion est qu'une approche locale ne devrait pas uniquement se baser sur la maximisation de la corrélation locale. En effet, une image n'est composée que de signaux discontinus, de contours ou de régions qui ont une dimension finie dans l'image. Toutes les irrégularités des deux images ainsi que les dissimilitudes qui les séparent, engendrent un comportement du coefficient de corrélation qui est lui-même discontinu, en fonction de l'environnement du pixel considéré. L'effet produit est un mauvais traitement au niveau de la synthèse des coefficients de détails à injecter. Les résultats obtenus grâce à cet outil exploratoire corroborent mes conclusions, à savoir qu'il semble plus judicieux de partir d'une méthode de fusion globale, et de venir localement la perturber avec une approche locale.

Trois nouvelles méthodes de fusion ont été développées et sont présentées dans la partie suivante. Elles pallient au moins partiellement les défauts présentés jusqu'alors. Une analyse visuelle et statistique d'images fusionnées par ces méthodes et par des méthodes existantes a été établie et les résultats sont résumés par la suite. Pour cette présentation et par souci de simplicité, nous allons nous placer dans le cas où le *MSM* permet directement de passer de *res0* à *res1* puis à *res2* ... comme dans le cas de la présentation des diverses implantations d'*ARSIS*.

6.2 Description et évaluation de nouvelles méthodes de fusion

Ces nouvelles implantations sont disponibles dans l'atelier logiciel fusion (annexe 1).

6.2.1 ATWT-M3-FONC

La méthode de fusion de Ballester *et al.* (2003) a attiré notre attention grâce à la bonne qualité géométrique de ses produits de fusion. Elle nous a inspiré la méthode hybride suivante. Les produits provenant de *ATWT-M3* laissant apparaître un flou général sur l'image, nous avons eu l'idée d'associer aux produits de fusion provenant de *ATWT-M3* une minimisation du terme géométrique de la fonctionnelle d'énergie. Nous cherchons à rehausser les contours un peu trop progressifs, tout en évitant les problèmes de distorsion radiométrique globale et locale liés à l'imprécision du terme radiométrique et à l'implantation du terme d'attache aux données. Cette nouvelle méthode de fusion, nommée *ATWT-M3-FONC*, devrait respecter la propriété de cohérence grâce à la première méthode de fusion de type *ARSIS* tout en tentant de limiter son flou.

6.2.2 ATWT-SharpendedM3

Cette nouvelle méthode de fusion tente de tirer profit de l'homogénéité du modèle global *ATWT-M3*, tout en venant à nouveau localement rehausser les structures jugées trop floues. La variance injectée étant insuffisante dans *ATWT-M3*, nous avons cherché à combler au moins partiellement le manque d'énergie des hautes fréquences spatiales en amplifiant les coefficients d'ondelettes correspondants. Le coefficient de pondération des coefficients d'ondelettes est une fonction qui dépend de l'activité locale de l'image que nous avons définie comme une fonction des écart-types locaux.

Notons que cette méthode peut être combinée avec le *MSM GLP* pour donner un *GLP-SharpendedM3*.

6.2.2.1 IMM de la méthode de fusion ATWT-SharpendedM3

Les paramètres a_{M3} et b_{M3} du modèle global *M3* sont estimés à basse résolution spatiale (par un ajustement par les moindres carrés ou l'axe d'inertie), c'est-à-dire entre les plans de coefficients d'ondelettes Pan et MS situés entre *res1* et *res2*.

Soit cc le coefficient de corrélation local entre ces deux plans, défini pour le ratio 4 sur une fenêtre $t = 21$ pour l'algorithme *ATWT* comme utilisé dans l'implantation pratique de l'algorithme *ATWT-RWM* décrit au chapitre précédent. La taille de la fenêtre devient $t = 9$ pour l'algorithme *GLP* comme définie dans la description de *GLP-CBD*. Sur cette même fenêtre, on calcule également les paramètres η et $\sigma RelLoc(D_{A,1-2})$.

Soit $\sigma loc(D_{A,1-2})$ l'écart-type calculé sur la fenêtre pour le plan de détails de la modalité Pan $D_{A,1-2}$. Soit $\sigma glob(D_{A,1-2})$ l'écart-type calculé sur l'ensemble de la modalité Pan. On définit l'activité locale relative du plan de détails $D_{A,1-2}$, $\sigma RelLoc(D_{A,1-2})$ par :

$$\sigma RelLoc(D_{A,1-2}) = \sigma loc(D_{A,1-2}) / \sigma glob(D_{A,1-2}) \quad \text{Eq. 6.1}$$

De même, on calcule l'activité locale relative du plan d'ondelettes de la $k^{ième}$ modalité MS $\sigma RelLoc(D_{Bk,1-2})$. On définit la quantité β par :

$$\beta = [\sigma RelLoc(D_{Bk,1-2}) / \sigma RelLoc(D_{A,1-2})]^2 \quad \text{Eq. 6.2}$$

et si $\beta < 1$ alors $\beta = 1$.

Le coefficient η permet de pondérer et d'amplifier les coefficients d'ondelettes haute résolution des modalités MS calculés avec a_{M3} et b_{M3} . Il est défini de la manière suivante :

$$\begin{aligned} &\text{si } |cc| < 0.8, \text{ alors } \eta = 1 \\ &\text{sinon } \eta = 1 + \beta(|cc| - 0.8) \\ &\text{enfin, si } \eta \geq 2 \text{ alors } \eta = 2. \end{aligned} \quad \text{Eq. 6.3}$$

6.2.2.2 HRIMM de ATWT-SharpenedM3

Le HRIMM reçoit de l'IMM les coefficients a_{M3} et b_{M3} globaux, ainsi que les paramètres η et $\sigma_{RelLoc}(D_{A,1-2})$, et doit synthétiser le plan d'ondelettes $D_{Bk,0-1}$. On mesure l'évolution de l'activité locale dans les échelles pour l'image Pan grâce à l'expression suivante :

$$\gamma = \sigma_{RelLoc}(D_{A,1-2}) / \sigma_{RelLoc}(D_{A,0-1}) \quad \text{Eq. 6.4}$$

avec une contrainte $1 \leq \gamma \leq 2$. La taille de la fenêtre d'estimation des écart-types locaux doit s'adapter aux dimensions des structures étudiées ; nous l'avons empiriquement fixée à $t = 11$ pour ATWT et $t = 5$ pour GLP.

Le plan d'ondelettes MS situé entre $res0$ et $res1$ est donné par la relation :

$$D_{Bk,0-1} = \gamma \eta (a_{M3} D_{A,0-1} + b_{M3}) \quad \text{Eq. 6.5}$$

6.2.3 ATWT-EnhancedRWM

Le point de départ de cette nouvelle méthode de fusion est l'IMM-HRIMM RWM. Elle peut s'adapter à tous les MSM disponibles sur notre plate-forme IDL, à savoir ATWT, GLP et MallatDaub4. L'objectif de cette nouvelle méthode est d'éviter le faïencage en appliquant les paramètres globaux a_{M3} et b_{M3} lorsque la corrélation locale est insuffisante. Ainsi, on évite le suréchantillonnage des modalités MS lorsque la corrélation entre le Pan et les MS est localement trop faible. Afin de ne pas produire de discontinuité entre le modèle global et local, on utilise une fonction linéaire entre les deux modèles pondérés par la valeur absolue du coefficient de corrélation local, estimé à basse résolution spatiale entre les plans de détails Pan et MS entre $res1$ et $res2$.

6.2.3.1 IMM de la méthode de fusion ATWT-EnhancedRWM

Comme pour la méthode de fusion précédente, l'IMM calcule les paramètres a_{M3} et b_{M3} issus du modèle global M3, ainsi qu'une image de coefficients de corrélation locaux. De plus, on calcule les coefficients a_{RWM} et b_{RWM} locaux selon le modèle RWM. La taille de la fenêtre d'estimation est fixée pour le ratio 4 comme pour la méthode de fusion précédente, à savoir $t = 21$ pour le MSM ATWT, et $t = 9$ pour GLP.

6.2.3.2 HRIMM de ATWT-EnhancedRWM

Le modèle HRIMM initial consiste en une combinaison linéaire entre les paramètres globaux et locaux, pondérée par une fonction du coefficient de corrélation local cc . Si μ est une fonction de la valeur absolue du coefficient de corrélation, alors le plan d'ondelettes $D_{Bk,0-1}$ se calcule de la manière suivante :

$$D_{Bk,0-1} = (1-\mu) (a_{M3} D_{A,0-1} + b_{M3}) + \mu (a_{RWM} D_{A,0-1} + b_{RWM}) \quad \text{Eq. 6.6}$$

La fonction μ est par exemple $\mu = \sqrt[4]{|cc|}$ ou $(cc)^{1/4}$.

On sait déjà qu'il y a un problème important lorsque le modèle local entre en contradiction avec le global. Par exemple un a_{M3} positif avec et un a_{RWM} négatif, produira, au pixel (i, j) concerné, une valeur aléatoire de $D_{Bk,0-1}$.

Notons que le produit qui donne le plan de détails $D_{Bk,0-1}$ est un produit terme à terme qui sous-entend que les images a et $D_{A,0-1}$ ont les mêmes dimensions. Si on utilise un MSM décimé tel que GLP, il est nécessaire de suréchantillonner l'image de paramètres a estimés à une échelle inférieure en utilisant par exemple le même outil GLP.

6.3 Evaluation de la qualité des nouvelles méthodes

Pour évaluer les nouvelles méthodes de fusion par rapport aux méthodes existantes, nous avons appliqué le protocole d'évaluation décrit dans le chapitre 2 de cette thèse.

6.3.1 Evaluation visuelle des produits de fusion à haute résolution

Pour cette évaluation visuelle, nous avons bénéficié d'une seconde campagne d'expérimentation. Elle a eu lieu à l'école des mines de Paris à Sophia-Antipolis, du 12 au 16 juin 2006. Le mode opératoire de cette seconde campagne a été placé en annexe (annexe 3).

Quatre méthodes de fusion ont été jugées par les expérimentateurs. Nous avons tout d'abord repris la méthode *ATWT-M3* qui donnait les résultats visuels les plus satisfaisants lors de la première campagne. Ballester *et al.* (2003) n'avaient jamais eu l'occasion de tester leur algorithme sur une panoplie importante d'images et de satellites, et cette campagne était l'occasion de comparer sa qualité à d'autres méthodes de fusion. Enfin, les deux nouvelles méthodes *ATWT-M3-FONC* et *ATWT-SharpendedM3* ont été retenues pour l'étude. La méthode de fusion *ATWT-EnhancedRWM* n'a pas été sélectionnée car elle a présenté des artefacts lors d'analyses préliminaires comme le montre la partie 6.3.1.3.

Les interprètes d'images ont étudié vingt-neuf extraits répartis sur sept images différentes. Ces images avaient été acquises par les satellites Ikonos et Quickbird ou provenaient des simulations CNES que nous avons déjà utilisées dans cette thèse. Le ratio de fusion est donc 4 pour toutes les images. Les produits fusionnés ont été présentés sous forme de compositions colorées en vraies couleurs, c'est-à-dire dont les composantes sont les modalités rouge, verte et bleue. Aucune image en fausses couleurs n'a été proposée aux interprètes contrairement à la campagne précédente qui avait conclu à leur manque d'intérêt dans ce cadre. Le tableau 6.1 rappelle les résolutions spatiales et le nombre d'extraits pour chaque site.

<i>Satellite ou simulations</i>	<i>Résolution des modalités MS originales</i>	<i>Résolution des images Pan</i>	<i>Ville</i>	<i>Nombre d'extraits</i>
Ikonos	4 m	1 m	Fredericton (Canada)	4
			Hasselt (Belgique)	5
			Toulouse	3
Quickbird	2.8 m	0.7 m	Fredericton (Canada)	3
			Starckville (USA)	8
			Toulouse	3
CNES	1.6 m	0.4 m	Toulouse	3

Tableau 6.1 : récapitulatif des images pour la seconde campagne d'expérimentation.

La redondance des sites (Fredericton et Toulouse) pour différents satellites a permis l'observation de l'incidence des différentes méthodes de fusion en fonction du capteur et de la résolution spatiale. Notons que nous avons synthétisé les modalités Pan des différents extraits CNES à partir des coefficients de contribution spectrale de chaque modalité MS par rapport au Pan de l'imagerie Quickbird (chapitre 3 de cette thèse, équation 3.2).

L'objectif de cette nouvelle campagne était l'évaluation qualitative à haute résolution spatiale de ces différentes méthodes, de procéder à un classement relatif et enfin de déterminer si ces méthodes sont acceptables pour un usage opérationnel des produits de fusion dans le cadre d'une exploitation opérationnelle au sein du ministère de la défense. Préalablement à la campagne, nous avions défini avec les interprètes les sites d'intérêt dans les images et les différents aspects qu'ils devaient évaluer. Nous avons classé ces différents aspects en quatre grandes catégories, qui concernaient leur

impression générale sur l'image, une critique de la géométrie, suivie par une critique de la radiométrie, pour aboutir finalement à une étude généralisée sur la qualité de l'image en terme de bruit, flou, piqué ...

La partie qui suit reprend les illustrations du début de chapitre afin de montrer que les nouvelles méthodes de fusion testées lors de la campagne ne reproduisent plus, ou dans une moindre mesure, les artefacts relevés lors de la première campagne. Nous y associons un bilan qualitatif par méthode de fusion. Ensuite, on montre les résultats visuels sur ces illustrations pour la méthode *ATWT-EnhancedRWM*, non retenue pour cette analyse. Enfin, nous présentons les résultats quantitatifs obtenus par les méthodes de fusion de la campagne 2, ainsi que par *ATWT-EnhancedRWM* et quelques autres méthodes existantes de type *ARSIS*.

6.3.1.1 Les méthodes de la campagne 2 vis-à-vis des artefacts

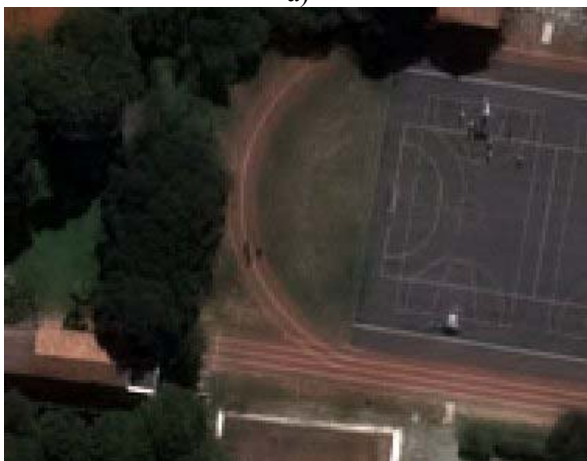
Les prochaines figures reprennent quatre des cinq illustrations présentées en début de chapitre. Nous y avons ajouté un exemple pour illustrer le défaut de l'inversion de contraste qui reste problématique pour la plupart des méthodes. Nous rappelons que les interprètes n'ont pas formé de mosaïque pour leur interprétation qualitative. Cependant, nous avons nous-même procédé à la création de ces mosaïques pour nous assurer que leurs conclusions demeuraient valables. Nous avons remarqué que les différences de couleurs n'étaient pas imputables à l'outil de visualisation. Les illustrations présentées ici ont toutes la même table de couleur au sein d'une même figure.



a)



b)



c)



d)



Figure 6.13 : image CNES, Toulouse. a) produit fusionné par *ATWT-M3*, b) par *FONC*, c) par *ATWT-M3-FONC*, d) par *ATWT-SharpendedM3*, e) modalités MS originales. Copyright CNES 2000.

e)

Aucune des quatre méthodes de fusion n'a produit de faïençage sur ce premier exemple. Les inversions de contraste et hétérogénéités de la piste d'athlétisme ne sont plus visibles. Il est dorénavant possible de suivre et donc de dessiner les traits des pistes de course et des différents terrains de sport. Le flou perceptible sur le ponctuel rectangulaire en haut à droite de la figure 6.13a donne une impression similaire dans l'image 6.13c, mais est nettement moins flou dans les deux derniers produits de fusion. La figure 6.13b présente une distorsion radiométrique globale qui tend vers le rose.



a)



b)

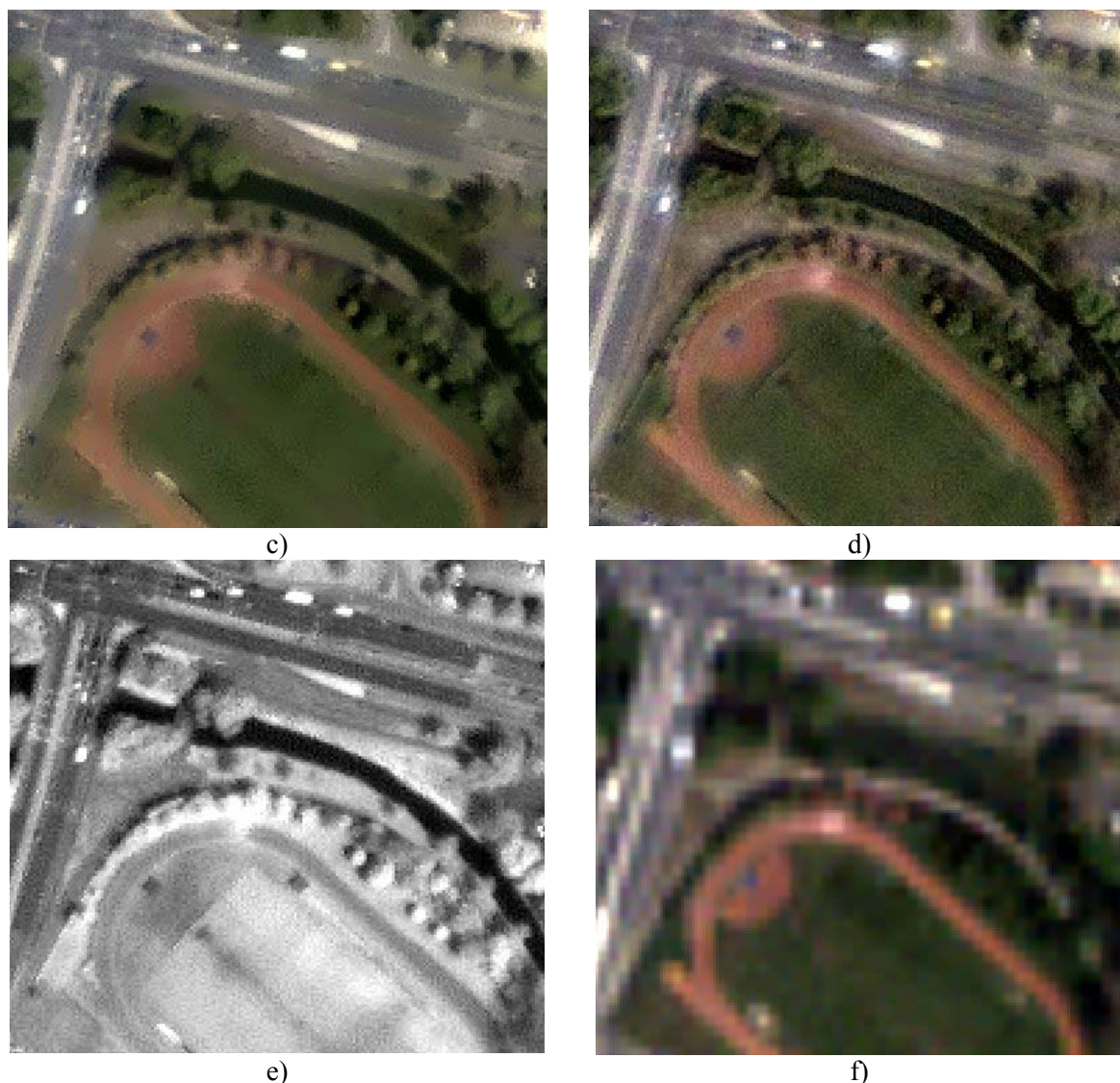


Figure 6.14 : image Ikonos, Hasselt (Belgique). a) produit fusionné par *ATWT-M3*, b) par *FONC*, c) par *ATWT-M3-FONC*, d) par *ATWT-SharpendedM3*, e) Pan, f) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Space Imaging 2002.

Les artefacts des méthodes de fusion de la première session étaient très nombreux sur cet exemple (figure 6.14). Ici, on remarque une fois encore l'absence de faïençage et une disparition des artefacts ponctuels colorés, comme la tache rose observée au bout de la piste du terrain de football. Par contre, des aplats de couleurs apparaissent pour les deux méthodes *FONC* et *ATWT-M3-FONC*. Par 'aplats', on entend des zones de l'image qui présentent à peu près la même couleur, comme avec des lignes de niveaux. Ce nouvel artefact est moins gênant mais reste un obstacle à une bonne interprétation de la composition colorée du produit fusionné.

On ne retrouve pas non plus le rebond de radiométrie plus élevé du milieu de la rivière pour toutes les méthodes. La radiométrie de la méthode *FONC* est mauvaise, avec des couleurs qui tendent vers le jaune. Les hautes radiométries saturées sont présentes tout en étant moins prononcées. On remarque que le rectangle blanc correspondant probablement à un bus ou un véhicule de transport de marchandises en haut au milieu de l'image est parfaitement synthétisé par la méthode *ATWT-M3-FONC*, sans aucun bavé sur la droite correspondant au déplacement du véhicule entre les différentes acquisitions Pan et MS. En effet, le flou à droite du véhicule donne l'impression qu'il a « allumé ses phares », mais il s'agit d'un phénomène que l'on appelle « le fantôme ». La méthode de fusion *ATWT-M3-FONC* a été fidèle au contenu géométrique de la modalité Pan, en ne tenant pas compte de l'information MS.

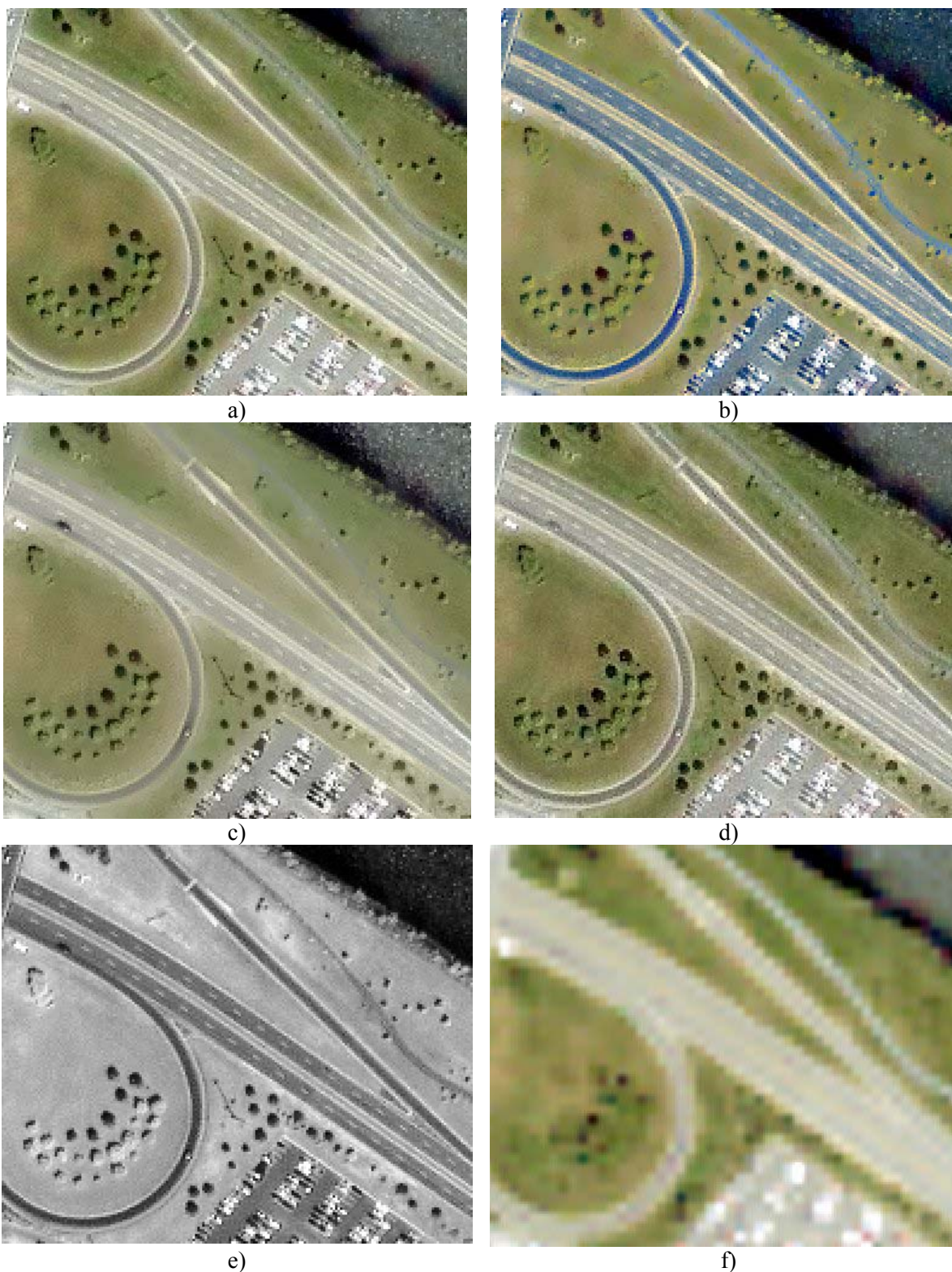


Figure 6.15 : image Ikonos, abords de Fredericton (Canada). a) produit fusionné par *ATWT-M3*, b) par *FONC*, c) par *ATWT-M3-FONC*, d) par *ATWT-sharpenedM3*, e) Pan, f) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Space Imaging 2002.

La troisième illustration (figure 6.15) avait montré lors de la première campagne un défaut d'inversion de contraste produit par les méthodes de fusion qui suivaient une approche locale. Ce défaut est désormais absent des quatre produits fusionnés présentés. La méthode de fusion *FONC* présente une forte distorsion des couleurs : le bâti (routes, bâtiments et véhicules) ainsi que la rivière

ne sont plus gris comme sur la composition colorée des modalités MS d'origine, mais sont teintés de bleu. Il s'agit du défaut majeur de cette méthode. Par contre, le rendu géométrique est très bon en ce qui concerne cette méthode et il en est de même pour la méthode *ATWT-SharpenedM3*, si on la compare avec la méthode *ATWT-M3* jugée un peu trop lisse sur cet exemple. Les aplats de couleurs déjà abordés dans l'exemple précédent (figure 6.14) sont bien visibles pour la méthode de fusion *ATWT-M3-FONC* placée figure 6.15c ; les transitions entre la route et la végétation ne sont plus très bien définies, ce qui rend difficile le tracé et la mesure des distances des différents éléments de l'image, et peut être dommageable pour l'interprétation.



a)



b)



c)



d)



e)



f)

Figure 6.16 : image Quickbird, centre ville de Fredericton (Canada). a) produit fusionné par *ATWT-M3*, b) par *FONC*, c) par *ATWT-M3-FONC*, d) par *ATWT-sharpenedM3*, e) Pan, f) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Digital Globe 2002.

En ce qui concerne le quatrième exemple (figure 6.16), la transition entre le bâtiment et son ombre avait été jugée trop molle pour *ATWT-M3*, ce qui donnait l'impression d'un dégradé entre les deux régions. On remarque ici la supériorité de la méthode *FONC* sur toutes les autres méthodes de fusion. Le rendu de la transition est parfait, très proche de celui de la modalité Pan en terme de qualité géométrique. Même si la couleur de la composition colorée de cet extrait tire vers le jaune, la qualité de synthèse des scènes situées en ville est très bonne. Les deux autres méthodes *ATWT-M3-FONC* et *ATWT-sharpenedM3* présentent une fidélité des couleurs vis-à-vis des modalités MS originales. Cependant, la transition bave un peu, particulièrement sur la figure 6.16c. La voiture rouge/orangée en haut à droite sur la route n'a pas été correctement synthétisée dans ces images ; soit la couleur se trouve légèrement décalée par rapport au véhicule, soit elle est diffusée dans la radiométrie des pixels aux alentours du bitume.

Nous avons sélectionné un exemple supplémentaire (figure 6.17) pour montrer le problème d'inversion de contraste. Cet exemple concerne aussi le stade de la ville de Toulouse, mais provient du satellite Ikonos. On remarque que la résolution spatiale de la modalité Pan (1 m) ne permet plus la distinction des différents couloirs de la piste d'athlétisme. Les méthodes de fusion ont toutes échoué dans la synthèse de la piste d'athlétisme, exceptée la méthode *FONC*.



a)



b)

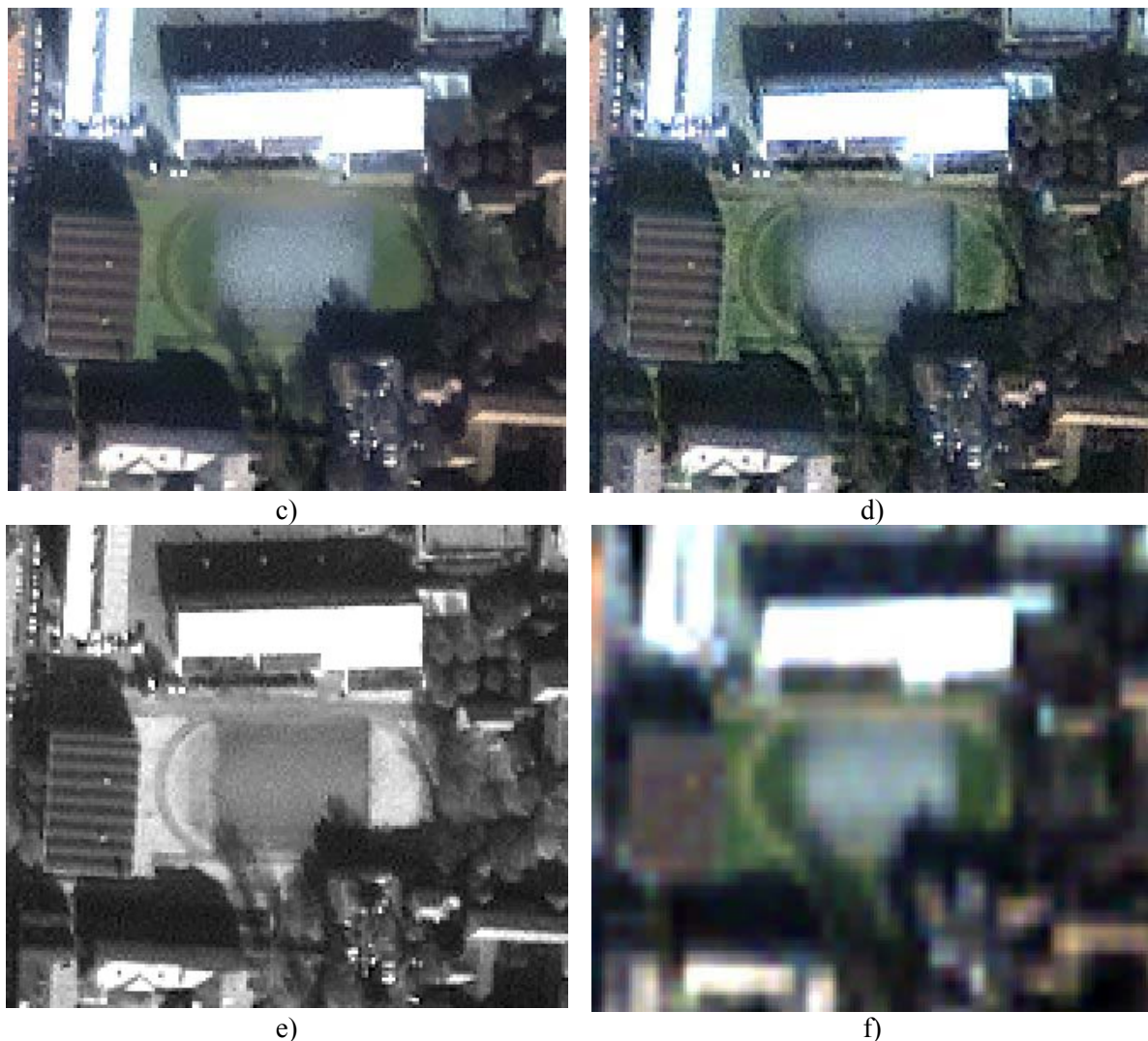


Figure 6.17 : image Ikonos, Toulouse. a) produit fusionné par *ATWT-M3*, b) par *FONC*, c) par *ATWT-M3-FONC*, d) par *ATWT-sharpenedM3*, e) Pan, f) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Space Imaging 2002.

En effet, les deux figures 6.17e et f montrent que la piste d'athlétisme est en inversion de contraste sur les deux images. Lors de la synthèse à haute résolution spatiale des modalités MS, la piste devrait apparaître plus claire que son environnement, ce qui n'est pas le cas des images 6.17a, c et d. On se retrouve tout à fait dans la situation où les approches locales se trouvent favorisées puisqu'elles ont la capacité de capter ces dissimilitudes avec une taille de fenêtre adaptée. Par contre, les approches globales se trouvent en échec puisque la corrélation globale entre la modalité Pan et chaque modalité MS est positive, ce qui engendre des coefficients de détails avec le même gradient que la modalité Pan. Ainsi, la méthode *FONC*, qui est la méthode se rapprochant le plus d'une approche locale, a été en mesure de capter cette dissimilitude et de synthétiser cet objet avec le bon gradient.

6.3.1.2 Synthèse de l'évaluation de la qualité par méthode de fusion

Méthode *ATWT-M3* (image a des figures 6.13 à 6.17) : sur les exemples testés lors de cette seconde campagne d'interprétation, les interprètes ont confirmé un bon comportement au niveau de la restitution des couleurs (aucune aberration chromatique). La végétation apparaît trop floue, et les hautes radiométries ponctuelles se trouvent souvent entourées d'un halo.

Méthode *FONC* (image b des figures 6.13 à 6.17) : les points forts de cette méthode sont une bonne restitution du contraste et un bon comportement de la méthode dans les zones artificielles

(conservation des géométries, ponctuels, linéaires et textures). Cette méthode est aussi capable de capter les dissimilitudes entre le Pan et les modalités MS. Cependant, l'inconvénient principal de cette méthode de fusion est la distorsion radiométrique, évidente pour les satellites Quickbird et surtout Ikonos. L'origine principale de cette distorsion est le terme radiométrique qui fait appel à des coefficients de pondération de la relation linéaire des modalités MS par rapport au Pan des méthodes de type contribution spectrale relative. Comme cette relation n'est qu'une approximation globale, elle n'est pas vérifiée précisément en chaque pixel de l'image, contrairement aux simulations CNES où la modalité Pan a été synthétisée comme telle.

Néanmoins, on note de très nombreux artefacts dans la végétation, avec des aplats qui dénaturent le paysage ; les textures sont atténuées et l'on perd les transitions entre la végétation et le bâti. Les défauts touchent aussi la couleur, avec la présence de halos, d'irisations colorées et des couleurs erronées qui ne permettent plus de distinguer les différents sols, qui donnent une impression non naturelle à l'image et faussent l'interprétation.

La perte des transitions se manifeste par des aplats, par un manque de netteté qui s'accompagne de la perte de ponctuels le long des transitions. Ce défaut apparaît aussi dans les extraits de la méthode *ATWT-M3-FONC*. La figure 6.18a montre un extrait fusionné par *FONC* où la présence des aplats est flagrante. Il n'est pas possible de distinguer le bord du chemin par rapport à la végétation. Les images b et c de la figure 6.18 correspondent au même extrait respectivement dans la modalité Pan et la composition colorée MS d'origine suréchantillonnée par un bicubique.

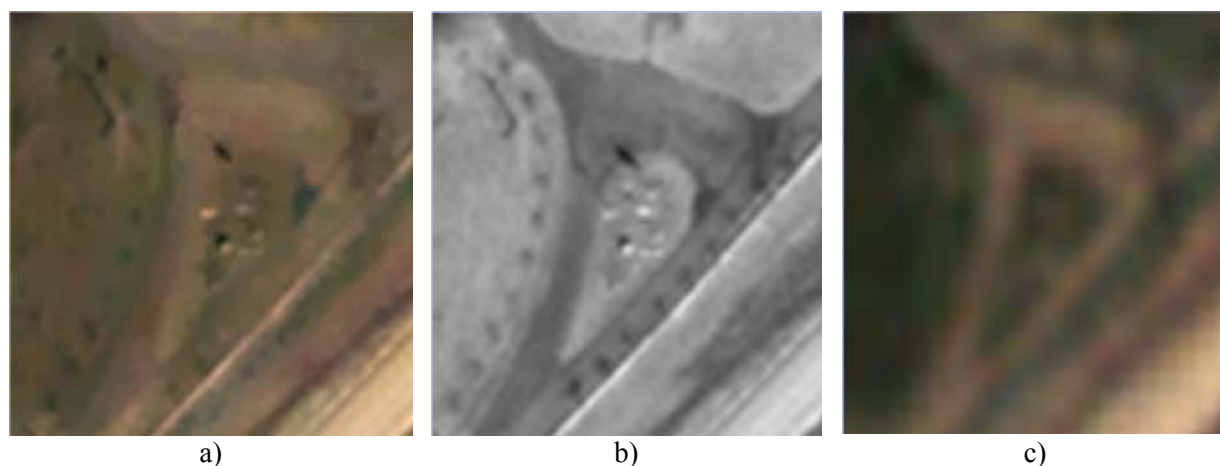


Figure 6.18 : image Quickbird, extrait de la périphérie de Starckville (USA). a) composition colorée obtenue par la méthode *FONC*, b) extrait Pan correspondant, c) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Digital Globe 2003.

La figure 6.19 suivante met l'accent sur le problème de distorsion locale engendrée par cette méthode de fusion. Comme pour l'illustration précédente, nous avons placé les modalités Pan et MS de référence figure 6.19b et c. Nous avons sélectionné un exemple de l'imagerie CNES pour montrer qu'il s'agit d'un problème indépendant de la distorsion radiométrique globale engendrée par le terme radiométrique. Cet extrait montre un immeuble orienté selon la diagonale ; on y voit très nettement des couleurs aberrantes apparaître sur le toit du bâtiment avec des dominantes violette et jaune. Les formes de haute radiométrie présentent un rebond violet parallèle et décalé de quelques pixels vers le bas et la droite de l'image. Quelques points du toit apparaissent aussi très nettement jaunes ou verts clairs. Il s'agit d'une mauvaise estimation de la radiométrie des pixels qui provient d'un, ou des deux autres termes de la méthode de Ballester *et al.* (2003).

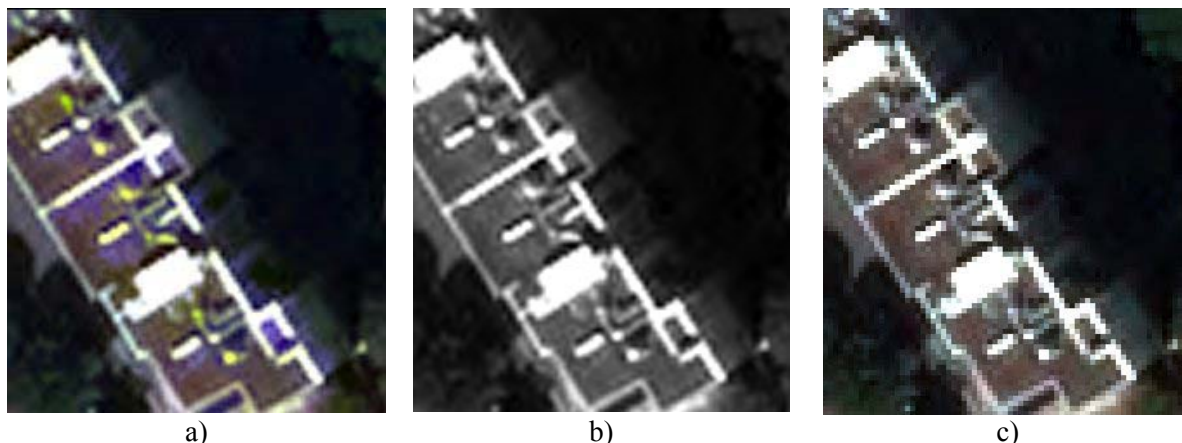


Figure 6.19 : image CNES, Toulouse. a) composition colorée obtenue par la méthode *FONC*, b) extrait Pan correspondant, c) modalités MS de référence à haute résolution spatiale (40 cm). Copyright CNES 2000.

Le non-respect de l'information colorée et le mauvais comportement de cet algorithme sur les linéaires, les réseaux, les surfaces et les contours ont poussé les interprètes à ne pas accepter cette méthode.

Méthode *ATWT-M3-FONC* (image c des figures 6.13 à 6.17) : les zones uniquement artificielles sont bien restituées par la méthode ; les géométries, les ponctuels et les linéaires ont été préservés. Le contraste est bon, dès lors que la végétation est absente de la scène. Le problème se corse nettement en présence de végétation, où les aplats de couleurs sont très nombreux. La méthode de fusion rencontre des difficultés à synthétiser les transitions artificiel/naturel, où le flou généré entraîne une perte des détails (ponctuels et linéaires). Il est possible d'y rencontrer des inversions de radiométrie. Les aplats rendent les couleurs un peu ternes. Cette méthode n'a pas été systématiquement rejetée par les interprètes, mais ils restent méfiants et mesurés dans leurs conclusions.

La figure 6.20a suivante est un exemple intéressant, puisqu'il illustre à la fois le point fort et le point faible de la méthode. Cet extrait montre un bâtiment dont la transition avec son ombre est particulièrement bien synthétisée, alors que la transition opposée, avec la végétation, est beaucoup plus floue. Les images placées en b et c représentent une fois encore les deux références Pan et MS suréchantillonnées correspondant à l'extrait d'intérêt.

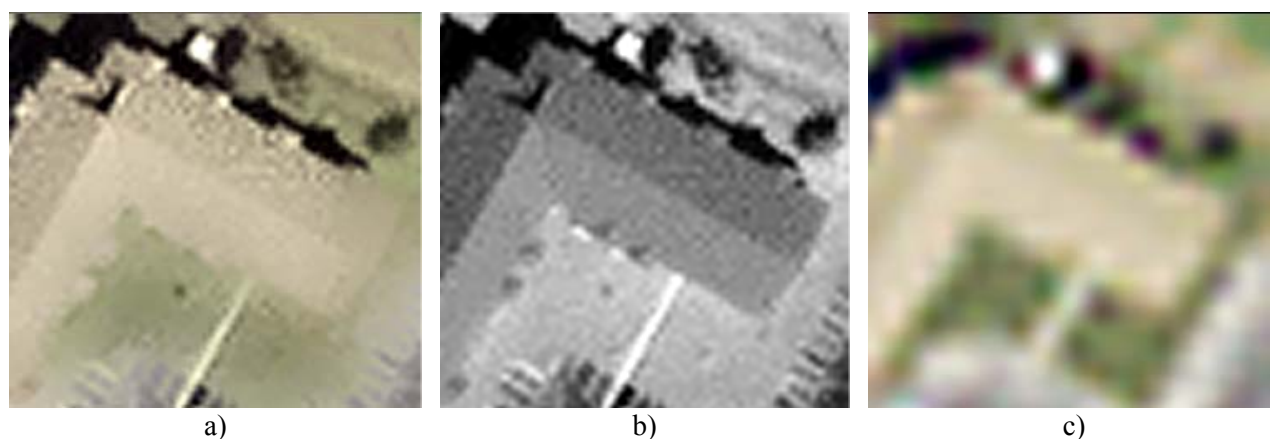


Figure 6.20 : image Ikonos, Fredericton (USA). a) composition colorée obtenue par la méthode *FONC*, b) extrait Pan correspondant, c) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Space Imaging 2002.

Le point commun entre ces deux méthodes est le terme géométrique. On rappelle que le terme géométrique prend la valeur du gradient des modalités MS, mais en tenant compte de la direction du gradient de la modalité Pan. La conclusion est que l'approximation numérique du gradient ne permet pas de retranscrire correctement les inversions de contraste.

Méthode *ATWT-SharpnedM3* (image d des figures 6.13 à 6.17) : cette méthode de fusion synthétise correctement les signatures spectrales puisqu'aucune aberration chromatique n'a été observée sur les exemples traités. Le comportement de la méthode vis-à-vis des textures sur la végétation a été jugé satisfaisant par les interprètes. On note l'apparition ponctuelle de flou et de halo d'une part, et d'autre part un ajout de bruit pour les images à plus haute résolution spatiale, c'est-à-dire pour l'imagerie Pléiades. Ces artefacts correspondent à un problème de compromis entre des coefficients d'ondelettes qui ont été trop ou pas assez rehaussés, donc un mauvais dosage dans l'injection des hautes fréquences spatiales.

Cette méthode de fusion a été sélectionnée comme celle présentant la meilleure qualité visuelle de cette deuxième session d'évaluation qualitative. Elle dépasse la qualité visuelle de la méthode *ATWT-M3* et a été désignée comme étant la seule à pouvoir être utilisée dans un cadre opérationnel sans la visualisation parallèle des modalités Pan et MS.

6.3.1.3 *Evaluation visuelle de la méthode ATWT-EnhancedRWM*

Les résultats préliminaires de cette méthode de fusion étaient très prometteurs, comme en témoignent les deux extraits de la figure 6.21. Ils correspondent aux extraits des figures 6.13 et 6.16.



Figure 6.21 : a) Composition colorée qui reprend l'exemple CNES au-dessus de Toulouse, issue de la fusion par *ATWT-EnhancedRWM*, b) extrait Quickbird au-dessus de Fredericton, même méthode de fusion.

Toutefois, en appliquant la méthode à d'autres extraits de la campagne 2, nous avons observé la présence d'artefacts. La figure 6.22 montre les deux extraits des illustrations 6.14 et 6.15 fusionnés par *ATWT-EnhancedRWM*. L'image de la ville de Hasselt avec le stade de foot montre des zones de flou ainsi que des amas de pixels colorés aberrants (rose). La route de l'image de droite arbore toutes les couleurs de l'arc-en-ciel. C'est très joli, néanmoins ce n'est pas l'objectif de la synthèse d'images à haute résolution spatiale.

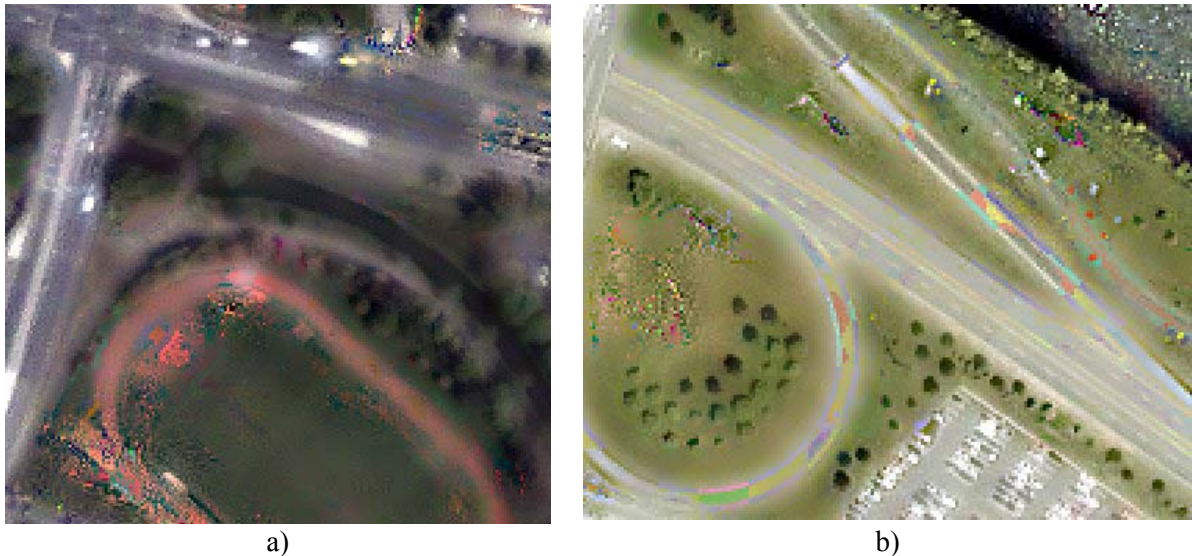


Figure 6.22 : a) Composition colorée qui reprend l'exemple Ikonos au-dessus de Hasselt (Belgique), issue de la fusion par *ATWT-EnhancedRWM*, b) extrait Ikonos au-dessus d'un échangeur de Fredericton (Canada), même méthode de fusion.

Ces deux extraits justifient la mise à l'écart de cette méthode de fusion dans le cadre de l'évaluation visuelle des produits de fusion pour la seconde campagne d'expérimentation.

6.3.2 Evaluation des propriétés de cohérence et de synthèse

Les étapes suivantes du protocole d'évaluation ont été appliquées aux images de la seconde campagne. Pour montrer un exemple de résultat obtenu grâce à la plate-forme IDL, nous avons choisi l'extrait de l'image Pléiades qui contient la piste d'athlétisme, utilisé pour nos illustrations. En plus des méthodes de fusion de la seconde campagne, nous y avons ajouté les produits de fusion des méthodes existantes *ATWT-RWM* et *ATWT-CBD* ainsi que ceux issus de *ATWT-EnhancedRWM*.

La vérification des deux propriétés nécessite l'utilisation d'un filtre de dégradation, que nous recommandons extérieur à celui utilisé dans les méthodes de fusion. Comme le *MSM* est *ATWT* pour toutes les méthodes de fusion, nous avons utilisé *GLP* pour la vérification des deux propriétés. Les distances choisies pour la qualité appartiennent au bilan de qualité que nous recommandons. Cependant, les deux distances de ce dernier, qui concernent la qualité géométrique, n'ont pu être calculées par manque de temps.

6.3.2.1 Vérification de la propriété de cohérence

La vérification de la propriété de cohérence correspond à la seconde étape du protocole d'évaluation de la qualité. Pour cela, tous les produits de fusion ont été dégradés jusqu'à la résolution spatiale des modalités MS d'origine : 160 cm pour les modalités CNES, 2.8 m pour Quickbird et 4 m pour Ikonos. Le tableau 6.2 donne un exemple de bilan de qualité, celui obtenu sur l'extrait CNES. Une remarque générale est que, quel que soit la méthode considérée, les résultats sont moins performants pour la modalité PIR que pour les autres modalités.

Discussion à propos des méthodes *ATWT-M3*, *ATWT-RWM*, *ATWT-CBD* et *ATWT-EnhancedRWM*: les valeurs des distances correspondant aux méthodes *ATWT-RWM* et *ATWT-EnhancedRWM* sont identiques pour les modalités visibles, et *ATWT-RWM* donne de meilleurs résultats pour la modalité PIR. Si on compare ces quatre méthodes, les modalités visibles ont été synthétisées au plus proche de leurs références par *ATWT-CBD* mais les résultats sont assez proches de ceux obtenus par *ATWT-RWM* et *ATWT-EnhancedRWM*. Par contre, la modalité PIR est mieux synthétisée par *ATWT-M3*.

Comparaison entre méthodes de la seconde campagne : en ce qui concerne les simulations CNES, on rappelle que la modalité Pan a été synthétisée à partir des quatre modalités MS. Les bandes

passantes des modalités MS couvrent ainsi parfaitement celle du Pan, donc certaines dissimilitudes ne vont pas exister avec ce type d'imagerie. Voilà pourquoi on s'attend aussi à ce que la méthode *FONC* offre un meilleur comportement avec les simulations CNES qu'avec les autres types d'imagerie.

Si on observe les résultats pour le biais, on remarque qu'il est proche de 0 pour toutes les méthodes de fusion et toutes les modalités, excepté pour les modalités rouge et verte synthétisées par *FONC*. Les valeurs de biais pour *FONC* traduisent la distorsion radiométrique rapportée lors de l'évaluation visuelle à basse résolution spatiale. Le biais est plus faible pour toutes les modalités MS synthétisées par *ATWT-M3*, avec une valeur inférieure ou égale à 0.2 %.

La différence en variance relative (référence - fusionnée) est la meilleure pour *ATWT-SharpendedM3* toutes modalités confondues. Cette différence est plutôt similaire pour *ATWT-M3* et *ATWT-M3-FONC* pour les modalités dans le visible. Par contre, sa valeur est doublée ou triplée pour *FONC* pour chaque modalité. Pour la modalité PIR, la différence de variance en valeur absolue de *ATWT-M3* vaut la moitié de celle obtenue par *ATWT-M3-FONC*, qui vaut elle-même la moitié de celle par *FONC*. Toutes les valeurs sont positives, sauf pour les modalités bleue et PIR de *FONC*, ce qui signifie que pour ces deux modalités, l'innovation a été trop importante tout en n'étant pas toujours pertinente qui peut alors être considérée comme un bruit.

L'écart-type de l'image de différence indique la proportion de pixels de l'image fusionnée en désaccord avec la référence. On remarque une fois encore la supériorité de *ATWT-SharpendedM3* sur les trois autres méthodes, à part cette fois-ci pour la modalité PIR. La valeur de l'écart-type relatif est moitié moindre pour cette méthode que pour *ATWT-M3* et *ATWT-M3-FONC*. *FONC* donne les valeurs d'écart-type les plus élevées. *ATWT-M3* est plus performante sur la modalité PIR. Les résultats de la méthode *ATWT-SharpendedM3* à propos de la modalité PIR nous pousse à prendre des précautions quant aux conclusions à apporter vis-à-vis de la méthode. Si elle donne les meilleurs résultats pour les modalités du visible en diminuant considérablement la différence en variance et l'écart-type, une exploitation conjointe de ces deux distances sur la modalité PIR montre que l'on a injecté une quantité plus importante de hautes fréquences spatiales (baisse de la différence de variance), mais comme parallèlement l'écart-type de l'image de différence a augmenté, on peut affirmer que l'information injectée est partiellement erronée.

Le coefficient de corrélation n'est pas très discriminant puisque sa valeur est souvent très proche de 1, sa valeur idéale.

Les distances multimodales choisies sont *ERGAS*, *SAM* et *biaisRel(diffNorms)*. Les valeurs sont très élevées pour *FONC* comparées aux autres méthodes. La meilleure valeur obtenue pour la distance *ERGAS* est celle trouvée pour *ATWT-SharpendedM3*. L'angle moyen entre les vecteurs spectraux des références et des produits fusionnés *SAM* est le plus faible pour *ATWT-M3*, mais reste assez proche de celui obtenu pour *ATWT-M3-FONC* et *ATWT-SharpendedM3*. Le biais relatif entre les normes des vecteurs spectraux est le meilleur pour *ATWT-M3*, et se dégrade progressivement en passant de *ATWT-SharpendedM3* à *ATWT-M3-FONC* puis *FONC*. Notons que l'angle des vecteurs MS et la différence de leurs normes présentent le défaut de ne pas prendre en compte la dynamique des images, comme le fait *ERGAS*. La dynamique de la modalité PIR est plus importante que pour les autres. Comme la qualité monomodale de la modalité PIR est moins bonne que pour les autres modalités, on obtient des valeurs moins bonnes pour les deux distances multimodales *SAM* et *biaisRel(diffNorms)*.

Conclusion sur la vérification de la propriété de cohérence pour toutes les méthodes de fusion et toutes les images de la campagne 2 : l'évaluation visuelle des produits de fusion ramenés à la basse résolution spatiale n'a pas révélé de distorsion flagrante. Néanmoins, les produits de fusion issus de la méthode *FONC* montrent des couleurs aberrantes même s'ils bénéficient d'une qualité géométrique satisfaisante. Au niveau de l'évaluation quantitative, la méthode de fusion *ATWT-SharpendedM3* est celle qui respecte le mieux la propriété de cohérence pour les modalités du visible. Pour ces modalités, ses résultats sont assez proches de ceux obtenus par *ATWT-CBD*, *ATWT-RWM* et *ATWT-*

EnhancedRWM. La modalité PIR a été synthétisée au plus proche de sa référence par *ATWT-SharpendedM3*, *ATWT-M3* et *ATWT-CBD*.

La distance *ERGAS* est inférieure à 1.5 pour toutes les images fusionnées, exceptée celle fusionnée grâce à la méthode *FONC*. Les deux meilleures méthodes à ce niveau sont *ATWT-SharpendedM3* et *ATWT-CBD* avec une valeur *ERGAS* inférieure à 1. Cette dernière méthode présente globalement la meilleure qualité multimodale pour tous les extraits.

En conclusion, les méthodes *ATWT-M3*, *ATWT-SharpendedM3* et *ATWT-CBD* respectent bien la propriété de cohérence. Les autres méthodes respectent plutôt bien cette propriété, mais la méthode *FONC*, quant à elle, ne la respecte que faiblement.

<i>Modalités</i>	<i>Distances</i>	<i>ATWT-M3</i>	<i>FONC</i>	<i>ATWT-M3-FONC</i>	<i>ATWT- SharpenedM3</i>	<i>ATWT- EnhancedRWM</i>	<i>ATWT-RWM</i>	<i>ATWT-CBD</i>
Bleue	<i>biaisRel</i>	0.1	0.4	0.6	0.3	0	0	0
	<i>diffVarRel</i>	6.5	-13.7	7.2	1.9	1.4	1.4	2.1
	<i>σRel</i>	4.4	16.5	4.4	2.8	2.7	2.7	2.4
	<i>cc</i>	0.997	0.993	0.997	0.999	0.999	0.999	0.999
Verte	<i>biaisRel</i>	0.1	13	0.6	0.4	0	0	0
	<i>diffVarRel</i>	6.7	21.8	7.4	1.8	1.5	1.5	2.1
	<i>σRel</i>	4.2	6.4	4.1	2.4	2.5	2.5	2.1
	<i>cc</i>	0.998	1.000	0.998	0.999	0.999	0.999	0.999
Rouge	<i>biaisRel</i>	0.1	12.7	0.5	0.4	0	0	0
	<i>diffVarRel</i>	5.9	18.1	6.4	0.9	0.9	0.9	1.8
	<i>σRel</i>	3.8	5.7	3.6	2.1	2.3	2.3	1.9
	<i>cc</i>	0.998	1.000	0.999	0.999	0.999	0.999	1.000
PIR	<i>biaisRel</i>	0.2	0.1	0.4	0.4	-1.3	-1.7	0
	<i>diffVarRel</i>	7.9	-24.6	13	2.2	-11.1	-20	8.4
	<i>σRel</i>	5.2	8.3	8.5	6.3	11.6	10	5.6
	<i>cc</i>	0.990	0.985	0.976	0.986	0.958	0.973	0.990
Distances multimodales	<i>ERGAS</i>	1.1	2.88	1.38	0.95	1.56	1.38	0.84
	<i>SAM</i>	2.33	6.25	2.94	3	2.8	2.81	1.4
	<i>biaisRel(diffNorms)</i>	2.8	23.3	6.5	4.9	-11.4	-16.5	0.3

Tableau 6.2 : résultats pour la vérification de la propriété de cohérence, extrait CNES

6.3.2.2 Vérification de la propriété de synthèse

Les ensembles originaux sont dégradés d'un ratio 4 par l'algorithme de décomposition hiérarchique de l'information *GLP*, avant d'appliquer toutes les méthodes de fusion à ces nouveaux ensembles. Il n'était pas nécessaire de procéder de la sorte pour les simulations CNES où les modalités MS à haute résolution (40 cm) étaient déjà disponibles. Les résultats qui concernent la vérification de la propriété de synthèse sur l'extrait CNES sont placés tableau 6.3.

Discussion à propos des méthodes *ATWT-M3*, *ATWT-RWM*, *ATWT-CBD* et *ATWT-EnhancedRWM* : concernant l'analyse monomodale, les mêmes remarques que précédemment peuvent être établies en comparant *ATWT-RWM* et *ATWT-EnhancedRWM*, à savoir qu'elles donnent les mêmes valeurs de distances pour les modalités du visible. La modalité PIR se révèle très pénalisante pour la qualité de *ATWT-EnhancedRWM* comme on peut le voir sur la figure 6.23. Toutes modalités confondues, les résultats délivrés par *ATWT-CBD* et *ATWT-RWM* sont bons et assez proches. Leurs résultats dépassent nettement ceux issus de la méthode *ATWT-M3*. Cette dernière montre la meilleure performance de ces quatre méthodes sur la modalité PIR, avec toutefois un aspect brumeux du résultat (figure 6.23a). Lors des analyses visuelles, *ATWT-M3* a été préférée à *ATWT-CBD* et *ATWT-RWM* comme nous l'avons discuté auparavant.



a)



b)



c)



d)

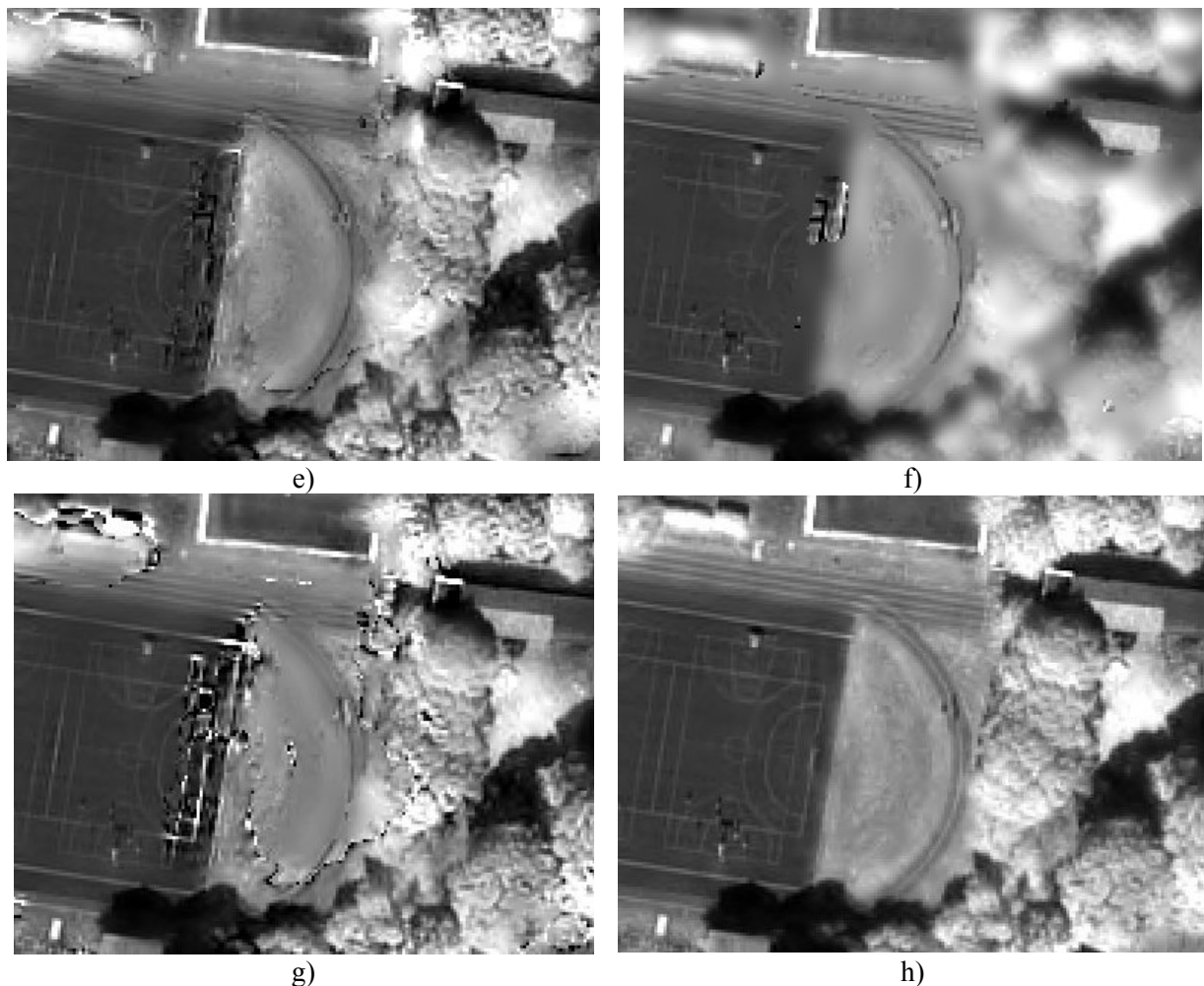


Figure 6.23 : image CNES, Toulouse. a) modalité PIR fusionnées par *ATWT-M3*, b) par *FONC*, c) par *ATWT-M3-FONC*, d) par *ATWT-SharpenedM3*, e) par *ATWT-RWM*, f) par *ATWT-CBD*, g) par *ATWT-EnhancedRWM*, h) modalité PIR originale. Copyright CNES 2000.

Comparaison entre méthodes de la seconde campagne : le biais obtenu lors de la vérification de la propriété de synthèse est très faible pour toutes les méthodes et toutes les modalités, excepté pour les modalités rouge et verte synthétisées par la méthode *FONC*. Ceci apparaît de manière évidente lors des analyses visuelles.

Comme nous l'avons expliqué précédemment, il est plus pertinent d'analyser la différence en variance et l'écart-type de l'image de différence en même temps pour pouvoir conclure. Les modalités bleue et PIR synthétisées par *FONC* montrent le même comportement au niveau de ces distances, c'est-à-dire une différence en variance négative et un écart-type élevé. Ces deux nombres caractérisent une injection trop importante de structures de haute fréquence équivalentes à un bruit. Au contraire, les modalités rouge et verte montrent conjointement à un écart-type élevé, une différence en variance élevée. Dans ce cas, l'injection a été insuffisante. *ATWT-SharpenedM3* présente les deux valeurs les plus faibles pour toutes les bandes spectrales du visible. Pourtant, les compositions colorées de certains extraits CNES ont donné une impression de bruit dans les zones homogènes de l'image. Les valeurs de ces deux distances du bilan de qualité monomodal sont moins bonnes dans le PIR. Une partie de l'énergie injectée n'est que du bruit. La figure 6.23d ne montre pas cet aspect bruité car l'extrait montre une zone riche en hautes fréquences spatiales, et montre au contraire un rehaussement des détails dans les arbres qui n'est pas aberrant.

La corrélation confirme les conclusions des deux distances précédentes. La méthode *FONC* donne la corrélation la plus faible par modalité. Les trois autres méthodes de la seconde campagne

montrent des coefficients de corrélation équivalents pour les modalités visibles. On observe une nette supériorité de *ATWT-M3* pour le PIR.

Les valeurs d'*ERGAS* sont proches et discréditent *FONC* par rapport aux trois autres méthodes. Les valeurs de la distance *SAM* favorisent *ATWT-M3* et *ATWT-M3-FONC*, alors que le biais relatif des normes est le plus faible pour *ATWT-M3*. Ainsi, les distances multimodales donnent l'avantage à *ATWT-M3* et *ATWT-SharpendedM3*.

<i>Modalités</i>	<i>Distances</i>	<i>ATWT-M3</i>	<i>FONC</i>	<i>ATWT-M3-FONC</i>	<i>ATWT-SharpendedM3</i>	<i>ATWT-EnhancedRWM</i>	<i>ATWT-RWM</i>	<i>ATWT-CBD</i>
Bleue	<i>biaisRel</i>	0.2	0.7	0.7	0.5	0.1	0.1	0.1
	<i>diffVarRel</i>	11.8	-6.7	12.1	3.8	5.9	5.9	5.2
	<i>σRel</i>	7.4	13.7	7.2	5	5.4	5.4	5
	<i>cc</i>	0.993	0.972	0.993	0.996	0.996	0.996	0.996
Verte	<i>biaisRel</i>	0.3	13.6	0.8	0.5	0.1	0.1	0.1
	<i>diffVarRel</i>	10.7	21.7	10.8	2.4	5	5	4.3
	<i>σRel</i>	6.4	7.1	6.1	4.1	4.6	4.6	3.9
	<i>cc</i>	0.995	0.999	0.996	0.997	0.997	0.997	0.998
Rouge	<i>biaisRel</i>	0.2	12.6	0.7	0.5	0.1	0.1	0.1
	<i>diffVarRel</i>	9	17.5	9.1	0.2	3.4	3.4	2.8
	<i>σRel</i>	5.4	6.3	4.9	3.9	4.1	4.1	3.1
	<i>cc</i>	0.997	0.985	0.998	0.998	0.998	0.998	0.999
PIR	<i>biaisRel</i>	0.2	0.3	0.4	0.4	-0.9	-1.1	0
	<i>diffVarRel</i>	4.4	-14.4	9.3	-6.7	-17.2	-10	11.7
	<i>σRel</i>	7.7	14.1	10.5	10	20.2	12.9	11.9
	<i>cc</i>	0.987	0.960	0.975	0.978	0.918	0.965	0.968
Distances multimodales	<i>ERGAS</i>	1.7	3.58	1.88	1.58	2.73	1.92	1.73
	<i>SAM</i>	3.25	6.44	3.26	8	5.77	3.51	2.73
	<i>biaisRel(diffNorms)</i>	2.5	27.1	6.2	4.4	-10.2	-11.2	0

Tableau 6.3 : résultat pour la vérification de la propriété de synthèse, extrait CNES

Conclusion sur la vérification de la propriété de synthèse pour toutes les méthodes de fusion et toutes les images de la campagne 2 : l'évaluation visuelle des produits de fusion des méthodes *ATWT-M3*, *ATWT-RWM* et *ATWT-CBD* a permis de constater les mêmes artefacts que lors de la première campagne d'analyse. En ce qui concerne les nouvelles méthodes et *FONC*, les conclusions de l'analyse visuelle des produits de fusion à haute résolution spatiale (partie 6.3.1) et celles de l'analyse visuelle pour la vérification de la propriété de synthèse correspondent. La méthode *FONC* montre une distorsion radiométrique générale des compositions colorées et la présence de quelques ponctuels de couleurs aberrantes. Les transitions entre le bâti et la végétation ne sont pas correctement synthétisées avec des aplats de couleurs qui atténuent le contraste qui peuvent faire disparaître des éléments ponctuels. Cette remarque a aussi été faite pour la méthode *ATWT-M3-FONC*. Par contre, ces méthodes montrent une très bonne performance dans la synthèse des structures urbaines telles que les bâtiments et les axes routiers. La méthode *ATWT-M3* confirme l'impression générale de flou dans les images. *ATWT-SharpendedM3* est celle qui montre la meilleure qualité visuelle, excepté sur les modalités CNES.

6.4 Conclusion

L'observation des artefacts de quelques méthodes de type *ARSIS* nous a apporté des pistes de développement pour de nouvelles méthodes de fusion. En effet, en dépit d'une bonne qualité de ces méthodes et particulièrement des méthodes qui adoptent une approche locale dans la synthèse haute résolution des coefficients de détails, les défauts observés affaiblissaient le bilan de qualité de ces méthodes. L'enjeu était de proposer de nouvelles méthodes de fusion qui conservaient la qualité quantitative de ces méthodes tout en luttant contre les artefacts identifiés.

Trois nouvelles méthodes ont donc été développées et évaluées en appliquant le protocole d'évaluation défini dans le chapitre 2 et en comparant ces résultats à ceux de méthodes existantes.

Par rapport à la référence qu'est *ATWT-M3*, on constate que la méthode *FONC* la dépasse en performance pour la qualité géométrique sur les zones urbaines, mais dégrade fortement les signatures spectrales des images. Ce non-respect de la propriété de cohérence est évité dans la méthode *ATWT-M3-FONC* qui utilise comme point de départ les produits fusionnés *ATWT-M3*. Par contre, si l'avantage de cette méthode par rapport à *ATWT-M3* est de mieux synthétiser les contours des bâtiments, on perd en qualité sur les régions naturelles. Par conséquent, grâce aux résultats statistiques, *ATWT-M3* reste globalement plus performante que ces deux méthodes.

De manière générale, c'est la méthode *ATWT-SharpendedM3* qui est la plus satisfaisante. Elle surpasse *ATWT-M3* en conservant de bons résultats pour la vérification de la propriété de cohérence et en offrant parallèlement une solution qui a permis de lutter efficacement contre le flou.

Toutefois, notons qu'il est difficile de conclure de manière catégorique quant aux performances d'une méthode relativement à d'autres. Comme l'interprétation des résultats l'a révélé, une méthode peut être satisfaisante sur un élément de l'image et pour une modalité donnée, et non pour d'autres. Ainsi, la méthode *ATWT-M3* a très souvent été surpassée en terme de qualité radiométrique et/ou géométrique par des méthodes de fusion qui n'ont pas été retenues lors de la seconde campagne ou parce qu'une ou plusieurs modalités donnaient de mauvais résultats en terme de distances. Des pistes d'amélioration de *ATWT-M3* ou *ATWT-SharpendedM3* peuvent être trouvées dans les autres méthodes *FONC* et *ATWT-EnhancedRWM*.

L'évaluation visuelle des produits de fusion à haute résolution spatiale en vraies couleurs par les interprètes d'images a mené à la conclusion que la méthode *ATWT-SharpendedM3* présente une qualité visuelle satisfaisante sur des images satellites (hors simulations), quasiment utilisable dans un contexte d'exploitation opérationnelle au sein du ministère de la défense.

Les interprètes émettent toutefois un bémol concernant le comportement de la méthode vis-à-vis des simulations CNES. En effet, la méthode n'est pas toujours en mesure de calculer le bon dosage

au niveau de l'injection de l'énergie dans les hautes fréquences spatiales. Certaines régions homogènes ont tendance à laisser apparaître un bruit correspondant à une injection trop importante de détails de hautes fréquences spatiales. Cette limite à propos de la méthode est confirmée par l'évaluation quantitative ; une diminution de la différence en variance relative conjointement à une augmentation de l'écart-type de l'image de différence comparée à la méthode *ATWT-M3* suggèrent une injection d'information haute fréquence erronée.

Dans la perspective de nouveaux développements, il faut noter que l'inversion de contraste est un problème qui risque d'être récurrent à cause des raisons exprimées dans la partie 6.1.3.5. L'enjeu du compromis entre approche locale et globale est toujours d'actualité. D'après la nouvelle méthode de fusion exploratoire développée avec Gabriel Peyré et placée en annexe (annexe 5), on peut conclure qu'une approche locale basée exclusivement sur la valeur du coefficient de corrélation locale ne permet pas de se prémunir d'un aspect hétérogène du produit fusionné final : on peut envisager que l'intervention d'une grandeur supplémentaire apporte une solution à cette limitation.

VII Conclusion, recommandations et perspectives

7.1 Conclusion

L'analyse de la bibliographie a montré la nécessité de définir très précisément le cadre de travail. La fusion d'images a de multiples applications, et choisir la synthèse de modalités spectrales MS à haute résolution spatiale engendre des contraintes qui ne sont pas toujours assimilées par la communauté. Il faut que les développeurs de nouvelles méthodes de fusion tiennent davantage compte des dissimilarités locales entre la modalité MS à fusionner et la modalité Pan à haute résolution spatiale qui contient les structures à injecter. Nous notons généralement une absence de considérations physiques dans l'établissement et l'évaluation des méthodes.

Il était essentiel de se pencher sur l'évaluation de la qualité avant de développer de nouvelles méthodes de fusion. Des lacunes ont été observées dans ce domaine et de nombreuses démarches maladroites ont été notées dans la littérature. Quelques travaux avaient toutefois tenté d'apporter un cadre à l'évaluation avec succès, tout en étant incomplets à notre avis. Nous avons apporté une généralisation de deux protocoles existants sous la forme d'un protocole à deux propriétés : cohérence et synthèse. Ce protocole est générique, capable d'être utilisé quel que soit le ratio, la résolution spatiale, le type de paysage considéré et le nombre de modalités.

Nous nous sommes ensuite interrogés sur l'obtention d'un bilan de qualité complet des produits de fusion. Nous avons étudié les distances utilisées par les développeurs et les utilisateurs d'images fusionnées. Une innovation de notre part a été la proposition d'une classification de ces distances en sept catégories et d'y associer la proposition d'un bilan de qualité qui compte une ou plusieurs distances de chaque catégorie visant à donner une idée complète de la qualité géométrique et radiométrique.

Nous avons cerné les limites de notre protocole. La première remarque est que le filtre de dégradation qui intervient dans le protocole a une influence sur la qualité. Notre expérience montre qu'il est préférable d'utiliser un filtre qui soit complètement différent de celui qui pourrait être utilisé dans les méthodes de fusion évaluées. Une autre limite est l'hypothèse d'extrapolation. En effet, l'évolution de la qualité des produits de fusion dans les échelles est un point controversé. Nous avons procédé à une étude empirique pour étudier les bornes de validité de cette hypothèse. Avec le mode opératoire que nous nous sommes fixés, nous concluons à une vérification fréquente de l'hypothèse d'extrapolation, en tenant compte du fait que la validation de cette hypothèse pour un cas donné dépend des bornes définies pour chaque distance du bilan de qualité choisi.

Nous avons porté un regard critique sur notre approche empirique de la vérification de l'hypothèse d'extrapolation. Une série de tests supplémentaires appliqués à des extraits des simulations CNES a permis de révéler un lien entre des bilans de qualité établis à des échelles de plus en plus fines et la quantité d'information à injecter lors de la fusion. Le peu d'images traitées n'a pas permis de définir des seuils capables de prédire le comportement des bilans de qualité dans les échelles, mais des pistes et observations ont été notées.

Nous avons contribué à l'enrichissement de notre bilan de qualité par une nouvelle distance pour l'évaluation de la qualité géométrique locale des produits de fusion. La fonction de transfert de

modulation des capteurs nous a inspiré cette distance car elle correspond à une courbe dans le domaine fréquentiel qui témoigne de la qualité en terme de bruit et de résolution des capteurs spatiaux. Comme cette grandeur est difficilement accessible, nous avons implanté nous-mêmes un module dans notre plate-forme qualité afin d'estimer cette information, qui s'appuie sur les contours contenus dans les images satellites. Ce module d'estimation a été validé avec des spécifications concernant l'imagerie Ikonos. La distance que nous proposons correspond à une norme $L2$ pour mesurer l'écart entre la courbe *FTM* obtenue pour un contour de référence et ce même contour fusionné. Les résultats obtenus avec cette distance sont en accord avec les analyses qualitative et quantitative obtenues pour des méthodes de fusion dont nous connaissions a priori leur incidence sur les contours.

Nous avons eu l'occasion de tester ce protocole et la plate-forme logicielle correspondante dans un cadre extérieur à la thèse. En effet, notre laboratoire a été choisi par le comité d'organisation du concours « fusion d'images » de IEEE pour évaluer les images fusionnées réalisées par huit méthodes de fusion sur des images de Quickbird et CNES. Nous avons pu ainsi vérifier que notre protocole s'applique à d'autres méthodes que celles que nous utilisons nous-mêmes. Le bilan de qualité que nous recommandions s'est avéré suffisant pour jauger la qualité des images fusionnées. Enfin, nous avons noté que le classement final proposé par le comité correspond à celui que l'on obtient avec la distance *ERGAS* seule. Par contre, en ce qui concerne cette dernière, nous n'avons pas trouvé de seuil général séparant les images fusionnées en bonne ou mauvaise qualité comme l'avance Wald (2002).

Munis de ces outils d'évaluation de la qualité, il nous était alors possible de proposer de nouvelles méthodes de fusion. Nous avons, tout d'abord, étudié les méthodes déjà publiées tout en cernant leurs avantages et inconvénients. Nous montrons que seules les méthodes de fusion qui font appel à un moment donné à l'analyse multiéchelle sont en mesure de respecter les modalités MS originales, et donc respectent la propriété de cohérence.

Outre le respect de la propriété de cohérence, les méthodes obéissant au concept *ARSIS* présentent généralement un bon respect de la propriété de synthèse, ce qui place quelques-unes des implantations du concept parmi les méthodes les plus performantes à ce jour. Cependant, la qualité visuelle à haute résolution spatiale est souvent jugée insuffisante. Nous avons identifié cinq catégories d'artefacts. Nous les illustrons par quelques exemples révélateurs et y apportons une explication en nous appuyant sur la théorie à l'origine du développement de la méthode de fusion et en reproduisant nous-même certains de ces défauts. La conclusion importante est que l'amélioration d'une méthode de fusion de type *ARSIS* correspond à un compromis entre une approche globale des relations inter-modalités (*IMM* et *HRIMM*) et une approche locale plus apte à traiter les dissimilitudes locales entre les modalités MS et Pan. Nous avons aussi relevé que le comportement des méthodes de fusion peut différer fortement en fonction de la modalité, et plus particulièrement en fonction de la corrélation entre le Pan et chaque MS.

Les nouvelles méthodes de fusion que nous proposons, explorent différentes stratégies de fusion qui favorisent plus ou moins l'approche globale par rapport à l'approche locale. Elles ont l'avantage de ne pas ou peu reproduire les artefacts relevés pour les précédentes méthodes pour les modalités du visible.

Parmi les nouvelles méthodes, la méthode de fusion appelée *ATWT-SharpendedM3* a été jugée comme étant celle présentant la meilleure qualité visuelle comparée aux autres implantations d'*ARSIS*. Elle présente également de très bons résultats concernant l'analyse quantitative de la qualité. Nous avons donc satisfait notre objectif, à savoir proposer une méthode plus performante.

D'un point de vue application, les méthodes de type *ARSIS* avaient été jugées avant cette thèse comme n'étant pas utilisables dans un cadre opérationnel d'organisation de mission de défense. Par contre, il est estimé que la nouvelle méthode *ATWT-SharpendedM3* est utilisable dans ce but.

Les analyses visuelles et bilans de qualité nous amènent à proposer l'idée de fusionner les modalités MS avec des méthodes de fusion différentes, comme par exemple fusionner les modalités visibles avec la méthode de fusion *ATWT-SharpnedM3* et la modalité PIR avec *ATWT-M3*. Des tests supplémentaires sont nécessaires pour définir un critère de sélection de la méthode de fusion, comme par exemple la corrélation globale entre la modalité Pan et chaque modalité MS.

Les nouvelles méthodes de fusion ne s'appliquent pas qu'aux modalités du visible, PIR et Pan puisqu'à aucun moment les spécificités de ces modalités ne rentrent en compte dans les modèles. Elles ne nécessitent pas un recouvrement total voire même partiel de la bande spectrale du Pan par celles des modalités MS. Ces méthodes peuvent être appliquées à l'infrarouge thermique (3 à 12 μm) ou aux micro-ondes (ondes de longueur d'onde centimétrique). Les modalités peuvent aussi être des résultats de modèles numériques présentés sous forme de grille rectangulaire. Par exemple, Cros (2004) aurait pu utiliser ces algorithmes pour fusionner ses cartes qui décrivaient les propriétés optiques des aérosols atmosphériques. La sortie du modèle donnait une grille en niveaux de gris dont les pixels les plus clairs correspondaient à l'épaisseur optique la plus importante. Ces modèles numériques peuvent aussi concerner des données climatiques ou géographiques comme par exemple l'orographie. Remund *et al.* (2003) ont montré que l'on pouvait utiliser les méthodes de fusion de type *ARSIS* pour cartographier la turbidité optique de l'atmosphère claire. Un autre domaine d'application est l'imagerie médicale. Elle fournit des images qui ont des caractéristiques différentes telles que des images fonctionnelles comme la tomographie par émission de positons et des images anatomiques de plus haute résolution spatiale comme les images à résonance magnétique. L'imagerie fonctionnelle ne contient généralement pas assez de détails anatomiques pour localiser précisément la position d'une tumeur ou d'autres lésions. La fusion entre ces deux types d'imagerie permet d'imposer une structure anatomique aux images fonctionnelles et donc d'améliorer et fiabiliser le diagnostic (Pajares et Cruz 2004). La figure 7.1 représente un résultat de ces auteurs.

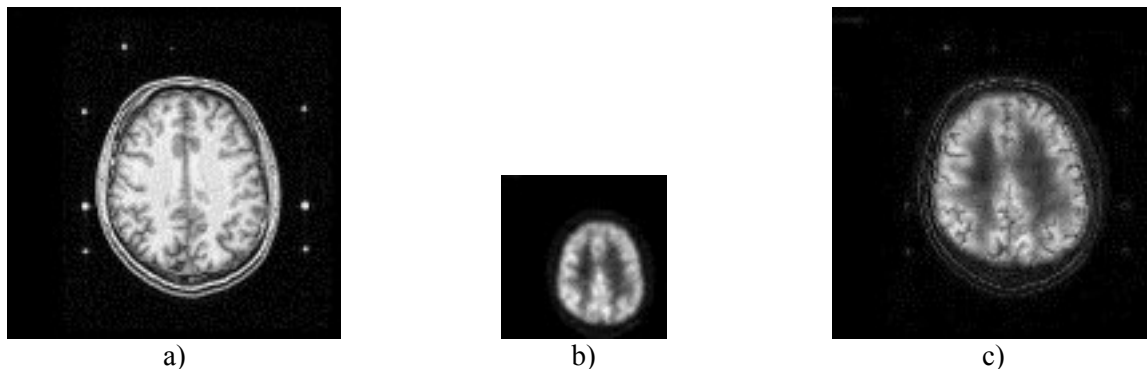


Figure 7.1 : a) image à résonance magnétique, b) tomographie par émission de positons; et c) produit fusionné à partir des images a et b. (Pajares et Cruz 2004)

7.2 Recommandations aux développeurs

La première recommandation peut paraître un peu évidente, mais nous préférons le souligner : il est nécessaire de bien cerner le cadre d'applications de la fusion dans lequel le développeur se place afin d'en saisir toutes les contraintes. Ensuite, les méthodes de fusion doivent être évaluées en appliquant le protocole d'évaluation de la qualité, sans qu'aucune étape ne soit négligée. Pour que les résultats issus des évaluations puissent être comparables, nous conseillons d'utiliser un filtre de dégradation qui soit le même pour la vérification des deux propriétés, mais indépendant des filtres utilisés dans les méthodes de fusion. Pour l'évaluation visuelle, les produits fusionnés en couleurs ou en niveaux de gris doivent être comparés à leurs références par l'intermédiaire de la création d'une mosaïque pour prévenir les distorsions radiométriques dues à l'outil de visualisation. L'utilisation de campagne d'expérimentation pour l'évaluation visuelle de la qualité des produits de fusion s'avère être un atout supplémentaire. Cependant, l'organisation de telles campagnes est très contraignantes au niveau de la logistique. Elle nécessite la venue pour quelques jours d'intervenants extérieurs, de mettre

à leur disposition une salle disposant de deux écrans par expérimentateur, ou paire d'expérimentateurs, de préférence des tubes cathodiques et non des écrans plats. De plus, les conclusions de l'évaluation visuelle dépendent aussi des domaines de travail des intervenants, et ne sont pas forcément généralisables à d'autres applications. En ce qui concerne l'analyse quantitative, il faut prendre garde à utiliser des distances qui soient complémentaires et non redondantes, et qui couvrent les différents aspects de la qualité des produits de fusion. En ce sens, le bilan de qualité proposé constitue une bonne base.

Nous conseillons aux développeurs d'avoir recours à l'analyse multiéchelle dans un des aspects de leurs nouvelles méthodes afin d'augmenter leur chance de vérifier la propriété de cohérence. D'après notre expérience, nous pouvons avancer que le concept *ARSIS*, dont les méthodes vérifient bien la propriété de cohérence, offre encore aujourd'hui un avenir prometteur de développement. Si le concept *ARSIS* est choisi pour développer de nouvelles méthodes, nous recommandons de privilégier des approches qui articulent judicieusement estimation globale et locale des paramètres des *IMM* et *HRIMM*.

7.3 Perspectives d'amélioration

7.3.1 Qualité

Un autre effort doit être fait dans la définition des trois distances multimodales suivantes : *SAM*, la différence des normes des vecteurs spectraux et la norme du vecteur résultant. Les limitations de ces distances multimodales sont leur dépendance vis-à-vis du nombre de modalités MS et de la dynamique de chacune d'entre elles. Elles ne suivent donc pas les recommandations à propos des distances multimodales de Wald (2000).

7.3.2 Régularité locale des images

Les images satellites sur lesquelles nous travaillons habituellement, présentent généralement des directions localement privilégiées dans l'image ; des routes qui s'intersectent, des bâtiments qui longent des axes routiers, les sillons de labour dans les champs... Si les images présentaient des structures isotropes et également distribuées sur l'image et de manière équivalente sur toutes les modalités, des approches globales prenant en compte toutes les modalités simultanément donneraient de bons résultats. Nous avons souligné à plusieurs reprises les dissimilitudes existantes entre les modalités Pan et chaque modalité MS ; cette idée sous-entend en théorie un traitement dissocié des bandes ainsi qu'un traitement dissocié des pixels, avec encore une fois le risque de créer des hétérogénéités dans l'image.

7.3.2.1 Classification des images MS à basse résolution spatiale

Une première idée est que l'on peut concevoir que des pixels proches dans l'image soient traités de manière différente s'ils n'appartiennent pas au même objet. Par contre, s'ils appartiennent au même objet, il semble normal qu'ils subissent à peu près le même traitement. La première idée que nous n'avons pas explorée, est de procéder à une segmentation préalable de l'image afin de regrouper les pixels et de leur associer un traitement différent par classe, en fonction de l'objet représenté. Par exemple, une méthode de fusion pourrait être appliquée sur le bâti et une autre sur la végétation. L'avantage serait d'assurer une homogénéité de la qualité visuelle par classe. Cependant, le comportement dans les pixels de transitions entre les différentes classes pourrait aussi créer des hétérogénéités. De plus, la qualité du résultat serait très dépendant de la méthode de classification utilisée.

7.3.2.2 MSM : bases de fonctions anisotropes

La régularité locale des images peut être captée par un *MSM* qui s'adapte au contenu de l'image contrairement à la transformée en ondelettes. Les ondelettes ont eu un immense succès ces vingt dernières années. En effet, elles ont été utilisées pour de nombreuses applications comme le filtrage, la déconvolution ou la compression d'images. Si les ondelettes sont particulièrement efficaces pour la détection de structures isotropes de différentes échelles, elles ne sont, par contre, pas optimales

pour la recherche d'objets anisotropes tels que des contours. De nouvelles transformées multiéchelles ont récemment été développées comme les curvelettes ou les bandelettes qui permettent de rechercher des objets de manière optimale, quand ces objets présentent de fortes anisotropies (Starck *et al.* 2003). Le point de départ de ces nouvelles transformées est souvent les plans d'ondelettes. Un post-traitement est appliqué aux coefficients d'ondelettes afin d'extraire des structures particulières, comme par exemple pour la transformée en curvelettes (Choi *et al.* 2005, Nencini *et al.* 2006).

Nous avons tenté de développer une méthode de fusion qui utilise la transformée en bandelettes. Nous ne rentrerons pas dans les détails de cette transformée. Les résultats concernant cette transformée n'ont pas été concluants, mais une idée importante peut être soulignée. La fusion concerne deux images : une image à la source des informations haute fréquence (Pan), et une image qui en est la cible (modalité MS). Par conséquent, si une base de fonctions est adaptée à chacune de ces deux images, comme ces images peuvent localement comporter des dissimilarités, leurs transformées sont aussi susceptibles d'être différentes. L'étape qui suit l'application du *MSM* pour les méthodes de type *ARSIS* est le calcul du modèle d'estimation des paramètres (*IMM*), qui ne peut pas être effectuée entre des images ayant été décomposées sur des bases de fonctions différentes. La conclusion de ces tests est que l'on a besoin d'une base commune de fonctions.

7.3.2.3 *IMM local*

L'*IMM* de la méthode exploratoire placée en annexe (annexe 5), que nous avons intitulée *MallatDaub79_WindowAdjustRWM*, correspond à une tentative de capter la régularité des paramètres de l'*IMM* dans la déformation des fenêtres d'estimation. Notons qu'il est possible d'utiliser le module développé sous IDL pour créer l'*IMM_WindowAdjustCBD* en remplaçant les plans de détails servant à l'estimation par les plans de contexte. Dans le passé, les *IMM* des couples *IMM-HRIMM RWM* et *CBD* utilisaient des fenêtres d'estimation de forme carrée qui ne sont pas adaptées à des objets singuliers.

Les tests préliminaires concernant cette méthode de fusion nous ont permis de montrer qu'une estimation des paramètres locaux de l'*IMM* avec déformation anisotrope de la fenêtre par maximisation de la corrélation locale est insuffisante pour s'affranchir des défauts d'hétérogénéités (faïençage) relevés lors de l'évaluation de la qualité visuelle. Des travaux supplémentaires doivent être effectués pour trouver le moyen d'ajuster le traitement des objets d'une modalité MS. Une déformation anisotrope de la fenêtre d'estimation pourrait être une solution, ce qui rejoindrait implicitement l'idée de la classification préalable des modalités MS originales. Le principe est de mesurer la ressemblance d'un pixel donné par rapport à ses pixels voisins en fonction de leur ressemblance, et non pas en fonction de leur distance par rapport pixel considéré. L'objectif serait d'attribuer à un pixel un comportement de fusion similaire à ceux de son entourage qui lui ressemblent.

7.3.2.4 *De l'IMM vers le HRIMM*

Dans le cas des méthodes de fusion qui utilisent un *MSM* décimé ainsi qu'une approche locale (*GLP*, *MallatDaub4*, *MallatDaub79* ... en association avec *RWM*, *CBD*, *EnhancedRWM*, *WindowAdjustRWM* ...), une image de paramètres locaux est obtenue dont la taille est plus petite que celle des modalités MS fusionnées. Ainsi, une étape de suréchantillonnage de l'image des paramètres est nécessaire dans le but de s'ajuster aux dimensions des modalités fusionnées. Plusieurs opérateurs de suréchantillonnage sont disponibles dans les logiciels de traitement d'images. Cependant, ces opérateurs ne tiennent pas compte du contenu des images lors de l'augmentation de la taille de l'image. Il pourrait être intéressant de suréchantillonner les paramètres en tenant compte de la direction privilégiée des objets dans les approximations de l'image Pan aux différentes échelles. Les paramètres estimés à basse résolution spatiale au niveau de l'*IMM* seraient transmis au *HRIMM* de manière anisotrope.

Bibliographie

- Acerbi-Junior F.W., J.G.P.W. Clevers et M.E. Schaepman, 2006. The assessment of multi-sensor image fusion using wavelet transforms for mapping the Brazilian Savanna, *International Journal of Applied Observation and Geoinformation*, Digital Object Identifier 10.1016/j.jag.2006.01.001, Elsevier (Eds.), Vol. 8, N°4 : 278-288.
- Aiazzi B., L. Alparone, S. Baronti et R. Carlà, 1997. Pyramid-based multi-sensor image data fusion, *In Proceedings of SPIE Conference "Wavelet Applications in Signal and Image Processing"*, vol. 3169 : 224-235.
- Aiazzi B., L. Alparone, S. Baronti et A. Garzelli, 2002. Context-Driven Fusion of High Spatial and Spectral Resolution Images Based on Oversampled Multiresolution Analysis, *IEEE Transactions on Geosciences and Remote Sensing*, vol. 40, N°10 : 2300-2312.
- Aiazzi B., L. Alparone, S. Baronti, A. Garzelli et M. Selva, 2002. An MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas, *In 2nd GRSS/ISPRS Joint Workshop on "Remote Sensing and Data Fusion over Urban Areas"*, IEEE Ed. (03EX646), 22-23 May 2003, Berlin, Allemagne, ISBN 0-7803-7719-2 : 90-94.
- Aiazzi B., S. Baronti, L. Santutti, M. Selva et L. Alparone, 2003. Impact of irreversible data compression on spectral distortion of hyper-spectral data, *Geoinformation for European-wide Integration*, Thomas Benes (Eds.), Rotterdam, Netherlands : 107-112
- Aiazzi B., L. Alparone, S. Baronti et M. Selva, 2006. MS + Pan image fusion by enhanced Gram-Schmidt spectral sharpening, *In Proceedings of the 26th EARSeL Symposium "New Strategies for European Remote Sensing"*, 29-31 May 2006, Varsovie, Pologne
- Alonso Reyes R., M. Melgar, S. Fernandez, C. Thomas, T. Ranchin et L. Wald, 2005a. Fusion of SPOT5 multispectral and Ikonos panchromatic images. *In Proceedings of the 24th EARSeL Symposium "New Strategies for European Remote Sensing"*, 25-27 May 2004, Dubrovnik, Croatia, Oluic (Eds.), Millpress, Rotterdam, ISBN 90 5966 003 X : 369-375.
- Alonso Reyes R., M.J. Gutierrez, S. Fernandez, C. Thomas, T. Ranchin et L. Wald, 2005b. Evaluation of the quality of Quickbird fused products. *In Proceedings of the 24th EARSeL Symposium "New Strategies for European Remote Sensing"*, 25-27 May 2004, Dubrovnik, Croatia, Oluic (Eds.), Millpress, Rotterdam, ISBN 90 5966 003 X : 343-351.
- Alparone L., V. Cappellini, L. Mortelli, B. Aiazzi, S. Baronti et R. Carlà, 1998. A pyramid-based approach to multisensor image data fusion, *Future Trends in Remote Sensing* : 419-426.
- Alparone L., S. Baronti, A. Garzelli et F. Nencini, 2004. A global quality measurement of pan-sharpened multispectral imagery, *IEEE Geoscience and Remote Sensing Letters*, vol. 1, N°4 : 313-317.
- Alparone L., L. Wald, J. Chanussot, L. M. Bruce et P. E. Gamba, 2006. IGARSS'06 Data Fusion Contest: Fusion of Panchromatic and Multispectral Images, *In IEEE IGARSS Symposium*, Denver, Colorado, USA, 31 July – 4 August 2006.
- Anonymous, 1986. *Guide des utilisateurs des données SPOT*, CNES et SPOT Image (Eds.), Toulouse, France, 3 volumes, révisé en janvier 1991.
- Blanc P., T. Blu, T. Ranchin, L. Wald et R. Aloisi, 1998. Using iterated rational filter banks within ARSIS concept for producing Landsat multispectral images, *International Journal in Remote Sensing*, vol. 19, N°12 : 2331-2343.

- Blanc P., L. Wald et T. Ranchin, 1998. Importance and effect of co-registration quality in an example of “ pixel to pixel ” fusion process. *In* Proceedings of the 2nd International Conference “ Fusion of Earth Data ”, T. Ranchin et L. Wald (Eds.), published by SEE/URISCA, Nice, France, Janvier 1998 : 67-74.
- Ballester C., V. Caselles, B. Rougé et J. Verdera, 2003a. Une méthode géométrique de fusion des images P+XS, *Rapport CNES*.
- Ballester C., V. Caselles, J. Verdera et B. Rougé, 2003b. A variational model for P+XS image fusion, *In* Proceedings of IEEE workshop on variational, geometric and level sets method in computer vision, Faudegars et Parios (Eds.), published in INRIA, Sophia-Antipolis, France, 12 octobre 2003.
- Ballester C., V. Caselles, L. Igual, J. Verdera et B. Rougé, 2006. A variational model for P+XS image fusion, *International Journal of Computer Vision*, Digital Object Identifier 10.1007/s11263-0066852-x, to appear.
- Bunttilov V. et T. Bretschneider, 2004. Assessment of the method-inherent distortions in wavelet fusion, *In* Proceedings of 25th ACRS'04 « Data processing », 22-26 novembre, Chiang Mai, Thaïlande : 239-244.
- Burke Hubbard B., 1995. *Ondes et Ondelettes, La Saga d'un Outil Mathématique*, Pour la Science, diffusion Belin, Paris, France, 235 p.
- Byeungwoo J. et D.A. Landgrebe, 1999. Decision fusion approach for multitemporal classification, *In* IEEE transactions on Geoscience and Remote Sensing, Digital Object Identifier 10.1109/36.763278, vol. 37, N°3, Part 1 : 1227-1233.
- Carper W.J., T.M. Lillesand et R.W. Kiefer, 1990. The use of Intensity-Hue-Saturation transformation for merging SPOT panchromatic and multispectral image data, *Photogrammetric Engineering and Remote Sensing*, vol. 56, N°4 : 459-467.
- Chavez P.S. et A.Y. Kwarteng, 1989. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principle component analysis, *Photogrammetric Engineering and Remote Sensing*, vol. 55, N°3 : 339-348.
- Chavez P.S., S.C. Sides et J.A. Anderson, 1991. Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT Panchromatic, *Photogrammetric Engineering and Remote Sensing*, vol. 57, N°3 : 265-303.
- Chen H. et P. K. Varshney, 2005. A human perception inspired quality metric for image fusion based on regional information, *Information Fusion*, Digital Object Identifier 10.1016/j.inffus.2005.10.001, Elsevier (Eds.), to appear.
- Chibani Y. et A. Houacine, 2002. The joint use of IHS transform and redundant wavelet decomposition for fusing multispectral and panchromatic images, *International Journal of Remote Sensing*, vol. 23 : 3821-3833.
- Choi M., R.Y. Kim M.R. Nam et H.O. Kim, 2005. Fusion of multispectral and panchromatic satellite images using the curvelet transform, *IEEE Geoscience and Remote Sensing Letters*, vol. 2, N°2 : 136-140.
- Choi M., 2006. A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter, *IEEE transactions on Geosciences and Remote Sensing*, vol. 44, N°6 : 1672-1682.
- Civco D.L., Y. Wang et J.A. Silander, 1995. Characterizing forest ecosystems in Connecticut by integrating Landsat TM and SPOT panchromatic data, *In* Proceedings of the ACSM/ASPRS Annual Convention, Charlotte, NC., 2 (Reston VA: ASPRS) : 216-224.
- Constantinos S.P., M.S. Pattichis et E. Micheli-Tzanakou, 2001. Medical imaging fusion applications : an overview, *In* Record of the Thirty-fifth conference on Signals, Systems and Computers, Pacific Grove, Californie, vol. 2 : 1263-1267
- Cornet Y., M. Binard et J.P. Donnay, 2004. Intérêt de la fusion d'images à haute résolution spatiale pour la classification de l'occupation du sol en milieu urbain, *Revue internationale de Géomatique : Imagerie satellite*, European Journal of GID and Spatial Analysis, Lavoisier (Eds.) : 417-439.
- Couloignier I., T. Ranchin, V.P. Valtonen et L. Wald, 1998. Benefit of the future SPOT5 and of data fusion to urban mapping, *International Journal of Remote Sensing*, vol. 19, N°8 : 1519-1532.

- Cros S., 2004. Création d'une climatologie du rayonnement solaire incident en ondes courtes à l'aide d'images satellitaires, Thèse de Doctorat en Energétique, Ecole des Mines de Paris, 13 septembre 2004, 157 pages.
- Daubechies I. 1988. *Orthonormal bases of compactly supported wavelets*. Communications on Pure and Applied Mathematics, vol. 41 : 906-909.
- De Boissezon H. et F. Laporterie, 2003. Evaluations thématiques et statistiques de cinq algorithmes de fusion P/XS sur des simulations d'images PLEIADES-HR, *SFPT*, vol. 1, N°169 : 83-99.
- Dial G., H. Bowen, F. Gerlach, J. Grodecki et R. Oleszczuk, 2003. IKONOS satellite, imagery, and products, *Remote Sensing of Environment*, vol. 88 : 23-36.
- Diemer C. et J. Hill, 2000. Local correlation approach for the fusion of remote sensing data with different spatial resolutions, *In Proceedings of the third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"*, 26-28 January 2000, publié par SEE/URISCA, Nice, Sophia-Antipolis, France, Ranchin T et Wald L. (Eds.): 167-182.
- Dutilleul P., 1989. An implementation of the « algorithmes à trous » to compute the wavelet transform. *In Combes, J.M., orossman, A., Tchamitchian, Ph. (Eds.), Wavelets: Time-Frequency Methods and Phase Space*. Springer, Berlin : 298-304.
- Ehlers P., 1991. Multi-sensor image fusion techniques in remote sensing, *ISPRS Journal Photogrammetry and Remote Sensing*, vol. 46, N°3 : 19-30.
- Eskicioglu A.M. et P.S. Fisher 1995. Image quality measure and their performance. *IEEE Transactions and Communications*, vol. 43, N°12 : 2959-2965.
- Filippidis A, L.C. Jain et N. Martin, 2000. Multisensor data fusion for surface land-mine detection, *IEEE Transactions on Systems, Man and Cybernetics, Part C*, vol. 30, N°1 : 145-150.
- Fung T. et E. LeDrew, 1987. Application on principal components analysis to change detection, *Photogrammetric Engineering and Remote Sensing*, vol. 53, N°12 : 1649-1658.
- Gamba P., J. Chanussot et L.M. Bruce, 2006. Contest organized by the Data Fusion Technical Committee at IGARSS 2006, *IEEE Geoscience and Remote Sensing Society*, December 2006. **NON CITE**
- Garzelli A., F. Nencini, L. Alparone, B. Aiazzi et S. Baronti, 2004. Pan-Sharpening of Multispectral Images : A Critical Review and Comparison, *In proceedings of IEEE Geoscience and Remote Sensing Symposium, IGARSS'04*.
- Garzelli A. et F. Nencini, 2005. Interband structure modeling for Pan-sharpening of very high resolution multispectral images, *International Journal on Information Fusion*, Digital Object Identifier 10.1016/j.inffus.2004.06.008, vol. 6, N°3 : 213-224. **CITE?**
- Garguet-Duport B., J. Girel, J.M. Chassery, et G. Pautou, 1996. The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data, *Photogrammetric Engineering and Remote Sensing*, vol. 62, N°9 : 1057-1066.
- Gillespie A.R., A.B. Kahle et R.E. Walker, 1987. Color enhancement of highly correlated images - II Channel ratio and 'chromacity' transformation techniques, *Remote Sensing Environment*, vol. 22 : 343-365.
- Gonzalez-Audicana M. et M. Seco, 2003. Fusion of multispectral and panchromatic images using wavelet transform. Evaluation of crop classification accuracy, *In Proceedings of the 22th EARSeL Symposium "Geoinformation for European-wide integration"*, 4-6 June 2002, Prague, Czech republic, Benes (Eds.), Millpress, Rotterdam, the Netherlands : 265-272.
- Gonzalez-Audicana M., J.L. Saleta, O.G. Catalan et R. Garcia, 2004. Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition, *IEEE Transactions on Geosciences and Remote Sensing*, vol. 42, N°6 : 1291-1299.

- Gonzalez-Audicana M., X. Otazu, O. Fors et J. Alvarez-Mozos, 2006. A low computational-cost method to fused Ikonos images using the spectral response function of its sensors, *IEEE Transactions on Geosciences and Remote Sensing*, vol. 44, N°6 : 1683-1691.
- Gonzalez R.C. et R.E. Woods, 1992. *Digital Image Processing*, Addison-Wesley, Reading, MA.
- Grégoire C., N. Dondainas et M. Raffy, 1993. Construction de modèles multispectraux pour milieux hétérogènes à partir de modèles validés sur milieux homogènes : application au LAI, *In proceedings of Deuxième réunion sur les changements d'échelle dans les modèles de l'environnement et de la télédétection*, Strasbourg, France, 17-19 mai 1993 : 5-11.
- Gungor O. et J. Shan, 2004. Evaluation of satellite image fusion using wavelet transform, *In proceedings of 20th Congress ISPRS "Geo-Imagery Bridging Continents"*, 12-13 July 2004, Istanbul, Turkey, vol. 35, part B7.
- Haydn R., G.W. Dalke, J. Henkel et J.E. Bare, 1982. Application of the IHS color transform to the processing of multisensor data and image enhancement, *In. Proceedings of International Symposium on Remote Sensing of Arid and Semi-Arid Lands*, Caire, Egypt : 599-616
- Hall D.L., 1992. *Mathematical techniques in multisensor data fusion* (Norwood: Artech House Inc.)
- Hendersen-Sellers A., M. Wilson et G. Thomas, 1985. The effect of spatial resolution on archives of land cover type, *Climatic Change*, vol. 7 : 391-399.
- Holst G.C., 1995. Electro-optical imaging system performance, Bellingham, WA: SPIE Optical Engineering Press : 146-147.
- Huang P.S. et T. Te-Ming, 2003. E target fusion-based approach for classifying high spatial resolution imagery, *In Workshops "Advances in Techniques for Analysis of Remotely Sensed Data"*, 27-28 Oct. 2003 : 175-181.
- IRARS, 1996. Civil NIIRS Reference Guide, Imagery Resolution Assessment and Reporting Standards Committee, U.S. Government, Washington, D.C., 4 p., 3 annexes.
- Kong X.N. et D. Vidal-Madjar, 1988. Effet de la résolution spatiale sur les propriétés statistiques des images satellites : une étude de cas, *International Journal of Remote Sensing*, vol. 9, 1315-1328.
- Kristjansson J.E., 1991. Cloud parametrization at different horizontal resolutions, *Quarterly Journal of the Royal Meteorology Society*, Vol. 117 : 1255-1280.
- Latry C, 2003. La chaîne de fusion de données SPOT5, *SFPT*, vol. 1, N°169 : 46-52.
- Lei F. et H. Tiziani, 1989. Modulation transfer function obtained from image structures, *In Proceedings and Workshops "High precision navigation"*, K. Linkwithz and U. Hangleiter (Eds.), Springer, Heidelberg : 366-377.
- Li H., 1994. Multi-sensor image fusion using the wavelet transform, *In ICIP-94.*, IEEE International conference, vol. 1 : 51-54.
- Li J., 2000. Spatial quality evaluation of fusion of different resolution images, *In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Amsterdam, Pays-Bas, vol. 33, B2-2 : 339-346.
- Li, S., Adelson, E.H., Wang, Y. 2001. Combination of images with diverse focuses using the spatial frequency. *InfusionsThérapie*, vol. 2, N°3 : 169-176.
- Lillesand T. M. et R. W. Kiefer, 1994. *Remote Sensing and Image Interpretation*. Third edition, John Wiley & Sons, New York, 750 p.
- Lillo-Saavedra, M., Gonzalo, C., Arquero, A., Martinez, E., 2005. Fusion of multispectral and panchromatic satellite imagery based on tailored filtering in the Fourier domain. *International Journal of Remote Sensing*, 26, N°6 : 1263-1268.
- Liu J.G. et J. McM. Moore, 1998. Pixel block mixing modulation: adding spatial detail to TM band 6 thermal imagery, *International Journal of Remote Sensing*, vol. 19 : 2477-2491.

- Liu J.G., 2000. Smoothing Filter-based Intensity Modulation: a spectral preserve image fusion technique for improving spatial details, *International Journal of Remote Sensing*, vol. 21, N°18: 3461-3472.
- Luxen M. et W. Förstner, 2002. Characterizing Image Quality: Blind Estimation of the Point Spread Function from a Single Image. In *Proceedings of Photogrammetric Computer Vision 2002*, ISPRS commission III, September 9-13, 2002, Gratz, Austria : 205-210.
- Mallat S., 1989. A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, N°7 : 674-693.
- Mallat S., 1998. *A wavelet tour of signal processing*, Academic press, USA, ISBN 0-12-466605-1, 577 p.
- Mangolini M., 1994. *Apport de la fusion d'images satellitaires multicapteurs au niveau pixel en télédétection et photo-interprétation*. Thèse de Doctorat en Sciences de l'Ingénieur, Université de Nice - Sophia Antipolis, Nice, France, 174 p.
- Meentemeyer V., 1987. Scale effects in landscape studies, *Ecological Studies*, vol. 64 : 15-21.
- Meyer Y., S. Jaffard et O. Rioul, 1987. L'analyse par ondelettes, *Pour la Science*.
- Meyer Y., 1992. *Les Ondelettes, Algorithmes et Applications*, Armand Colin (Eds.), Paris, France, 172 p.
- Munichika C. K., J. S. Warnick, C. Salvaggio et J. R. Schott, 1993. Resolution enhancement of multispectral image data to improve classification accuracy. *Photogrammetric Engineering & Remote Sensing*, vol. 59, N°1 : 67-72.
- Nencini F., A. Garzelli, S. Baronti et L. Alparone, 2006. Remote sensing image fusion using the curvelet transform, *Information Fusion*, Digital Object Identifier 10.1016/j.inffus.2006.02.001, Elsevier (Eds.), to appear.
- Nunez J., X. Otazu, O. Fors, A. Prades, V. Palà et R. Arbiol, 1999. Multiresolution-based image fusion with additive wavelet decomposition, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, N°3 : 1204-1211.
- Otazu X., M. Gonzalez-Audicana, O. Fors et J. Nunez, 2005. Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, N°10 : 2376-2385.
- Otazu X., Gonzalez-Audicana M. et O. Fors, 2006. A framework to introduce sensor spectral response into image fusion methods. The WiSpeR and FAST-WiSpeR implementations, In *Proceedings of the 26th EARSeL Symposium "New Strategies for European Remote Sensing"*, 29-31 May 2006, Varsovie, Pologne.
- Pajares G. et J.M. Cruz, 2004. A wavelet-based image fusion tutorial, *Pattern Recognition*, doi:10.1016/j.patcog.2004.03.010, vol. 37, N°9 : 1855-1872.
- Pellemans A, R. Jordans et R. Allewijn, 1993. Merging multispectral and panchromatic SPOT images with respect to the radiometric properties of the sensor, *Photogrammetric Engineering & Remote Sensing*, vol. 59, N°1 : 81-87.
- Piella G., 2003. A general framework for multiresolution image fusion : from pixels to regions, *Information Fusion*, vol. 4 : 259-280.
- Piella G. et H. Heijmans, 2003. A new quality metric for image fusion, In *Proceedings of the IEEE International conference on Image Processings*, vol. 3/4: 173-176.
- Pohl C., Y. Wang et B.N. Koopmans, 1994. The 1995 flood in the Netherland space, *ITC Journal*, vol. 4 : 414-415.
- Pohl C. et J.L. Van Genderen, 1998. Multisensor image fusion in remote sensing: concepts, methods and applications, *International Journal of Remote Sensing*, vol. 19, N°5 : 823-854.

- Pradines D., 1986. Improving SPOT image size and multispectral resolution. *In Proceedings SPIE Conference "Earth Remote Sensing using the Landsat Thematic Mapper and SPOT Systems"*, 660 : 78-102.
- Price J.C., 1999. Combining multispectral data of differing spatial resolution, *IEEE Transactions on Geosciences and Remote Sensing*, vol. 37, N°3 : 1199-1203.
- Profanter C., R. Prommegger, M. Gabriel, R. Moncayo, G. J. Wetscher, T. Lang et R. Bale, 2004. Computed axial tomography-MIBI image fusion for preoperative localization in primary hyperparathyroidism, *The American Journal of Surgery*, vol. 187, N°3 : 383-387.
- Puissant A. et J. Hirsch, 2004. Télédétection urbaine et résolution spatiale optimale : intérêt pour les utilisateurs et aide pour les classifications, *Revue internationale de Géomatique*, Lavoisier (Eds.) : 403-415.
- Raffy M., 1993. Remotely-sensed quantification of covered areas and spatial resolution. *International Journal of Remote Sensing*, vol. 14, N°1 : 135-159.
- Raffy M., 1994. The role of spatial resolution in quantification problems: spatialization method, *International Journal of Remote Sensing*, vol. 15, N°12 : 2381-2392.
- Ranchin T. et L. Wald, 1992. The wavelet transform for the analysis of remotely sensed images, *International Journal of Remote Sensing*, vol. 14, N°3 : 615-619.
- Ranchin T. et Wald, L. 2000. Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing*, vol. 66, N°1 : 4-18.
- Ranchin T., B. Aiazzi, L. Alparone, S. Baronti et L. Wald, 2003. Image fusion-the ARSIS concept and some successful implementation schemes, *International Journal of Remote Sensing*, Digital Object Identifier 10.1016/S0924-2716(03)00013-3, Elsevier (Eds.), vol. 58, N°1-2 : 4-18.
- Raptis V.S., Vaughan R.A., T. Ranchin et L. Wald, 1998. Assessment of different data fusion methods for the classification of an urban environment, *In Proceedings of the second conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"*, 28-30 January 1998, publié par SEE/URISCA, Nice, Sophia-Antipolis, France, Ranchin T et Wald L. (Eds.): 167-182.
- Redondo R., F. Sroubek, S. Fisher et G. Cristobal, 2005. Multifocus fusion with multisize windows, *Applications of Digital Image Processing XXVIII*, Andrew G. et Tescher (Eds.), Proc. SPIE, vol. 5909 : 410-418.
- Remund J., Wald L., Lefevre M., Ranchin T., Page J., 2003. Worldwide Linke turbidity information. *In proceedings of ISES Solar World Congress*, 16-19 June, Göteborg, Sweden, CD-ROM published by International Solar Energy Society.
- Ricchetti E., 2004. Application of optical high resolution satellite imagery for archaeological prospection over Hierapolis (Turkey), *In Proceedings of Geoscience and Remote Sensing Symposium, IGARSS'04*, IEEE International, Digital Object Identifier 10.1109/IGARSS.2004.1369976, vol. 6 : 3898-3901.
- Rokhmatuloh M.A., R. Tateishi, K. Wikantika et K. Munadi, 2003. Study on the spectral quality preservation derived from multisensor image fusion techniques between JERS-1 SAR and Landsat TM data, *In proceedings IEEE International Geoscience and Remote Sensing Symposium IGARSS 2003*, 21-25 July 2003, Toulouse, France.
- Ryan R., Baldridge B., Schowengerdt R.A., Choi T., Helder D.L. et S. Blonski, 2003. IKONOS spatial resolution and image interpretability characterization, *Remote Sensing of Environment*, vol. 88 : 37-52.
- Sengupta S.K., A.S. Lopez et D.N. Poland, 2004. Class-label statistics: a basis for fusing information from multispectral imagery with an application to unsupervised detection of human settlement, *In Proceedings of Geoscience and Remote Sensing Symposium, IGARSS'04*, IEEE International, Digital Object Identifier 10.1109/IGARSS.2004.1368600, vol. 2 : 1084-1087.
- Shettigara V.K., 1992. A generalized component substitution technique for spatial enhancement of multispectral images using higher resolution data set, *Photogrammetric Engineering and Remote Sensing*, vol. 58, N°5 : 561-567.

- Shi W., C. Zhu, Y. Tian et J. Nichol, 2005. Wavelet-based image fusion and quality assessment, *International Journal of Applied Earth Observation and Geoinformation*, vol. 6 : 241-251.
- Smith A.R., 1978. Colour gamut transform pairs, *Computer Graphics*, vol. 12 : 12-19.
- Starck J.L., F. Murtagh, E.J. Candès et D.L. Donoho, 2003. Gray and color image contrast enhancement by the curvelet transform, *IEEE Transactions on Image Processing*, vol. 12, N°6 : 706-716.
- Terretaz P., 1997. Comparison of different methods to merge SPOT P and XS data: Evaluation in an urban area. In Proceedings of the 17th Symposium of EARSeL, *Future Trends in Remote Sensing*, Gudmansen P. (Eds.), Lyngby, Denmark, 17-20 June. A.A. Balkema, Rotterdam : 435-445.
- Thomas C., Ranchin T., Wald L. et Lanéri, J.C., 2005. A modular platform for fusion of images. In Proceedings of the Fourth International Conference Physics in Image and Signal Processing PSIP'2005, January 31-February 2 2005, Toulouse, France, ISBN 2-912328-22-5 : 269-273.
- Thomas C. et Wald L., 2005. Assessment of the quality of fused products. In Proceedings of the 24th EARSeL Symposium "New Strategies for European Remote Sensing", 25-27 May 2004, Dubrovnik, Croatia, Oluic (Eds.), Millpress, Rotterdam, ISBN 90 5966 003 X : 317-325.
- Thomas C. et L. Wald, 2006a. Analysis of changes in quality assessment with scales, In Proceedings of FUSION06, 10-13 July 2006, Florence, Italie.
- Thomas C. et L. Wald, 2006b. A MTF-based distance for the assessment of the geometrical quality of fused products, In Proceedings of FUSION06, 10-13 July 2006, Florence, Italie.
- Thomas C. et Wald L., 2006c. Comparing distances for quality assessment of fused products, In Proceedings of the 26th EARSeL Symposium "New Strategies for European Remote Sensing", 29-31 May 2006, Varsovie, Pologne.
- Tsagaris V., V. Anastassopoulos et G.A. Lampropoulos, 2005. Fusion of hyperspectral data using segmented PCT for color representation and classification, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, N°10 : 2365-2375.
- Tu T.M., S.C. Su, H.C. Shyu and P.S. Huang, 2001. A new look at IHS-like image fusion methods, *Information Fusion*, vol. 2, N°3 : 177-186.
- Tu T.M., P.S. Huang, C.L. Hung et C.P. Chang, 2004. A fast intensity-hue-saturation technique with spectral adjustment for Ikonos imagery, *IEEE Geosciences and Remote-Sensing Letters*, vol. 1, N°4 : 309-312.
- Vijayaraj V.; C. O'Hara et N. Younan, 2004. Quality analysis of pan-sharpened images, In Proceedings of Geoscience and Remote Sensing Symposium, IGARSS'04, IEEE International, Digital Object Identifier 10.1109/IGARSS.2004.1368951, vol. 1 : 85-88.
- Vrabel J., 1996. Multispectral imagery band sharpening study, *Photogrammetric Engineering and Remote Sensing*, vol. 58, N°9 : 1075-1083.
- Vrabel J., 2000. Multispectral imagery advanced band sharpening study, *Photogrammetric Engineering and Remote Sensing*, vol. 66, N°1 : 73-79.
- Wald L. et T. Ranchin, 1995. Fusion of images and raster-maps of different spatial resolutions by encrustation: an improved approach. *Computers, Environment and Urban Systems*, vol. 19, N°2 : 77-87.
- Wald L., T. Ranchin et M. Mangolini, 1997. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images, *Photogrammetric Engineering and Remote Sensing*, vol. 63, N°6 : 691-699.
- Wald L., 1999. Some terms of reference in data fusion, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, N°3 : 1190-1193.
- Wald L., 2000. Quality of high resolution synthesised images: Is there a simple criterion ?. In Proceedings of the third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", Sophia Antipolis, France, January 26-28, 2000, Thierry Ranchin and Lucien Wald (Eds.), published by SEE/URISCA, Nice, France, pp 99-103.

- Wald L., 2002. *Data Fusion: Definitions and Architectures. Fusion of Images of Different Spatial Resolutions*, Les Presses de l'Ecole des Mines (Eds.), Paris, 197 p.
- Wang Z. et A.C. Bovik, 2002. A universal image quality index, *IEEE Signal Processing Lett., Transactions on Pattern Analysis and Machine Intelligence*, vol. 9 : 81-84.
- Wang Z., D. Ziou, C. Armenakis, D. Li et Q. Li, 2005. A comparative analysis of image fusion methods, *Transactions on Geoscience and Remote Sensing*, vol. 43, N°6 : 81-84.
- Welch R.M., M.S. Navar et S.K. Sengupta, 1989. The effect of spatial resolution upon texture-based cloud field classifications, *Journal of Geophysical Research*, vol. 94 (D.12) : 14767-14781.
- Woodcock C. E. et A. H. Strahler, 1987. The factor of scale in remote sensing. *Remote Sensing of Environment*, vol. 21 : 311-332.
- Xydeas C. et V. Petrovic, 2000. Objective image fusion performance measure, *Electronics Letters*, N°36 : 308-309.
- Yang W., F. Cauneau, J.P. Paris et T. Ranchin, 2000. Influence of landscape changes on the results of the fusion of P and XS images by different methods. In Proceedings of the third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images", 26-28 January 2000. SEE/URISCA, Nice, Sophia-Antipolis, France, Ranchin T. et Wald L. (Eds.): 47-56.
- Yang X.H., Z.L. Jing, G. Liu, L.-Z. Hua et D.W. Ma, 2006. Fusion of multi-spectral and panchromatic images using fuzzy rule, *Nonlinear Science and Numerical Simulation*, Digital Object Identifier 10.1016/j.cnsns.2005.12.001, Elsevier (Eds.), to appear.
- Yocky, D. A., 1996. Multiresolution wavelet decomposition image merger of Landsat Thematic Mapper and SPOT panchromatic data, *Photogrammetric Engineering and Remote Sensing*, vol. 62, N°9 : 1067-1074.
- Zhang Q. et J. Wang, 2002. Detection of buildings from Landsat-7 ETM+ and SPOT panchromatic data in Beijing, China, In Proceedings of Geoscience and Remote Sensing Symposium, IGARSS'02, IEEE International, Digital Object Identifier 10.1109/IGARSS.2002.1026841, vol. 5 : 2977-2979.
- Zhang Y., 1999. A new merging method and its spectral and spatial effects, *International Journal of Remote Sensing*, vol. 20, N°10 : 2003-2014.
- Zheng Y., E.A. Essock, B.C. Hansen et E.M. Haun, 2005. A new metric based on extended spatial frequency and its application to DWT based fusion algorithms, *Information Fusion*, Digital Object Identifier 10.1016/j.inffus.2005.04.003, Elsevier (Eds.), to appear.
- Zhijun W., D. Ziou et C. Armenakis, 2004. Production of high-resolution remote sensing images for navigation information infrastructures, *Geospatial information science*, vol. 7, N°2 : 129-134.
- Zhou J., D.L. Civco et J.A. Silvander, 1998. A wavelet transform method to merge Landsat TM and SPOT panchromatic data, *International Journal of Remote Sensing*, vol. 19, N°4 : 743-757.

Liste des figures

<i>Indice de la figure</i>	<i>Intitulé de la figure</i>	<i># page</i>
1.1	Image Ikonos, partie de l'aéroport (Blagnac) de Toulouse. a) Composition colorée initiale des modalités visibles, 4 m de résolution spatiale, suréchantillonnées par un opérateur bicubique, b) composition colorée des images fusionnées à 1 m de résolution. Copyright Space Imaging 2002.	4
1.2	Image SPOT2, barrage des Trois Gorges (Chine). a) Modalité Pan, acquise en 1990, b) modalité verte, 1998. Copyright CNES SPOT-Image 1990 et 1998.	5
1.3	Réponses normalisées des bandes spectrales du satellite Quickbird.	6
1.4	Image Quickbird, Fredericton (Canada) - illustration du phénomène d'occultation. a) Pan, 1 m, b) bleue, 4 m suréchantillonnée à 1 m (bicubique), c) extraits d'une ligne particulière de chaque image : le chemin d'herbe est absent dans l'image bleue. Copyright Digital Globe 2002.	7
1.5	Image SPOT2, Toulouse (France). a) modalité Pan, (10 m de résolution), b) modalité <i>XSI</i> , 20 m rééchantillonnée (bicubique). La chute d'eau n'apparaît pas sur cette modalité. Copyright CNES 1990.	8
1.6	Illustration du phénomène d'inversion de contraste – image Quickbird, Fredericton (Canada). a) Pan, b) bleue, c) extraits d'une ligne particulière : le bâtiment présente un contraste inversé par rapport à son environnement dans les deux images. Copyright Digital Globe 2002.	9
1.7	Définition du vecteur spectral. On définit par $B(i, j)$ un vecteur MS de référence, et $B^*(i, j)$ le vecteur MS correspondant fusionné.	13
2.1	Méthode de visualisation de deux images : création d'une image intermédiaire (ou mosaïque) qui contient les deux images à comparer.	23
2.2	a) Modalité Pan (Quickbird) - Starckville (USA), b) simulation sous IDL d'un bruit Gaussien. Copyright Digital Globe 2004.	29
2.3	a) Modalité Pan (Ikonos) – Fredericton (Canada), b) même image ayant subi une rotation de 180 degrés. Copyright Space Imaging 2002.	30
2.4	a) Parcelles agricoles acquises par la modalité Pan (Ikonos) – Hasselt (Belgique), b) même modalité dont tous les comptes numériques ont été multipliés par deux. Copyright Space Imaging 2002.	31

2.5	a) Corrélogramme entre Q et le coefficient de corrélation pour le ratio 2 (corrélation entre les deux séries : 0.999), b) idem pour le ratio 4 (corrélation entre les deux séries : 0.989)	32
<hr/>		
3.1	Diversité des images. a) satellite Ikonos – banlieue de Hasselt en Belgique, b) parcelles agricoles de la même zone géographique, c) échangeur autoroutier, Fredericton au Canada, Quickbird et d) centre ville de Marseille, SPOT5. Copyright respectifs pour Ikonos, Quickbird, et SPOT5 : Space Imaging 2002 , Digital Globe 2002, CNES 2000.	36
3.2	Image SPOT5, composition colorée PIR, rouge et verte – collines de la région marseillaise – Copyright CNES 2002.	38
3.3	Centre ville de Madrid (Espagne), composition colorée rouge, verte et bleue – Quickbird – Copyright Digital Globe 2000.	39
3.4	Spectre d'une image quelconque et résultats successifs de la transformée en ondelettes dyadique. Le spectre de Fourier représenté en fonction des fréquences normalisées est séparée en deux à chaque application de la transformée.	48
3.5	Modalités Pan. a) SPOT5, Marseille (France) centre ville, b) idem, collines provençales, c) Quickbird Madrid (Espagne) centre ville, d) idem, végétation éparse. Copyright CNES (2000) et Copyright Digital Globe (2000).	50
3.6	Image SPOT5 - extrait de la mer Méditerranée, zone près de Marseille. Copyright CNES 2000.	51
<hr/>		
4.1	Tache d'Airy. a) vue avec une élévation de 45° , b) vue de dessus.	59
4.2	TF d'un signal d'entrée $E(f)$ échantillonné (courbe rouge). La courbe rouge témoigne du repliement spectral qui correspond à la somme des courbes noires, signal de l'image de départ qui a été périodisé par l'échantillonnage tous les (kfe) avec $k \in \mathbb{Z}$.	60
4.4	a) Contour parfait, b) contour perçu à travers un système d'acquisition.	61
4.5	Principe de l'approche basée contour.	62
4.6	Image Ikonos – arête d'un bâtiment de Fredericton (Canada). a) a) contour extrait manuellement, b) moyennage du contour par un filtre passe-bas, c) Dérivation numérique par un opérateur de Sobel à une dimension, d) résultat de la dérivation sans filtrage passe-bas préalable. Copyright Space Imaging 2002.	62
4.7	Principe de la transformée de Hough, a) du domaine de l'image, vers b) et c) celui de Hough.	63

4.8	a) Contour reconstruit après le seuillage dans l'espace de Hough, b) contour reconstruit après seuillage de l'image a.	64
4.9	a) Rappel du contour traité, b) en trait plein, le contour suréchantillonné et moyenné. En pointillés, la sigmoïde correspondant le mieux au profil.	66
4.10	Résultat de la courbe obtenue pour le contour pour l'illustration.	67
4.11	Simulation Pléiades sur la ville de Toulouse. a) contour de référence, b) contour synthétisé par <i>ATWT-RWM</i> , c) <i>ATWT-M3</i> , et d) contour obtenu par <i>MallatDaub4-M3</i> . Copyright CNES 2000.	70
4.12	a) Courbes <i>CFC</i> pour la référence (courbe noire), pour le contour fusionné par <i>ATWT-RWM</i> (courbe bleue), par <i>ATWT-M3</i> (courbe verte) et enfin par <i>MallatDaub4-M3</i> (courbe rouge), b) à droite, valeurs de la norme L_2CFC entre la référence et chacune des méthodes de fusion.	71
4.13	Extrait d'un des bâtiments de la prison de Toulouse. Copyright CNES 2000.	72
4.14	Contour de référence (en trait plein) et la fonction sigmoïde ajustée.	72
4.15	Ratio 2 - courbes <i>CFC</i> pour la référence (courbe noire) et les cinq contours fusionnés et valeurs de la norme L_2CFC entre la référence et chacune des méthodes de fusion.	73
4.16	Résultats de l'ajustement par une fonction sigmoïde pour les cinq méthodes de fusion différentes et minimisation obtenue pour la distance $L2$ la plus importante obtenue pour <i>GLP-CBD</i> .	74
4.17	Ratio 4 - courbes <i>CFC</i> pour la référence (courbe noire) et les cinq contours fusionnés et, à droite, valeurs de la norme $L2$ entre la référence et chacune des méthodes de fusion.	74
4.18	Ratio 4 - résultats de l'ajustement par une fonction sigmoïde pour les cinq méthodes de fusion différentes, et ajustement obtenu pour la distance $L2$ la plus importante obtenue pour <i>GLP-CBD</i> .	75
4.19	Ratio 4 – images MS à 160 cm obtenues par <i>GLP</i> - courbes <i>CFC</i> pour la référence (courbe noire) et les cinq contours fusionnés et valeurs de la norme L_2CFC entre la référence et chacune des méthodes de fusion.	76
4.20	Ratio 4 – MS simulées à basse résolution par <i>GLP</i> . résultats de l'ajustement par une fonction sigmoïde pour les cinq méthodes de fusion différentes, et ajustement obtenu pour la distance $L2$ la plus importante obtenue (<i>ATWT-RWM</i>).	76
<hr/>		
5.1	Le modèle de couleurs <i>HSV</i> de Smith.	80
5.2	Décomposition hiérarchique de l'information.	85

5.3	Les trois classes de modèles du concept <i>ARSIS</i> .	86
5.4	Schéma descriptif du concept <i>ARSIS</i> (d'après Wald 2000).	87
5.5	schéma récapitulatif du couple <i>IMM-HRIMM M1</i> .	89
5.6	Schéma récapitulatif des deux couples <i>M2</i> et <i>M3</i> .	90
5.7	Schéma récapitulatif du couple <i>RWM</i> .	91

6.1	Image Pléiades, Toulouse - a) produit fusionné par <i>ATWT-M3</i> , b) par <i>ATWT-RWM</i> , c) par <i>GLP-CBD</i> , d) par <i>GLP-M3</i> , e) Pan. Copyright CNES 2000.	105
6.2	Figure 6.2 : image Ikonos, Hasselt (Belgique) - a) produit fusionné par <i>ATWT-M3</i> , b) par <i>ATWT-RWM</i> , c) par <i>GLP-CBD</i> , d) par <i>GLP-M3</i> , e) MS originale. Copyright Space Imaging 2002.	106
6.3	Figure 6.3 : image Ikonos, centre ville de Fredericton (Canada) - a) produit fusionné par <i>ATWT-M3</i> , b) par <i>ATWT-RWM</i> , c) par <i>GLP-CBD</i> , d) par <i>GLP-M3</i> , e) MS originale. Copyright Space Imaging 2002.	108
6.4	Image Ikonos, abords de Fredericton (Canada) - a) produit fusionné par <i>ATWT-M3</i> , b) par <i>ATWT-RWM</i> , c) par <i>GLP-CBD</i> , d) par <i>GLP-M3</i> , e) MS originale. Copyright Space Imaging 2002.	109
6.5	Image Quickbird, centre ville de Fredericton (Canada) - a) produit fusionné par <i>ATWT-M3</i> , b) par <i>ATWT-RWM</i> , c) par <i>GLP-CBD</i> , d) par <i>GLP-M3</i> , e) MS originale. Copyright Digital Globe 2002.	110
6.6	Image Ikonos, échangeur autoroutier, Fredericton (Canada). a) image couleur fusionnée par <i>ATWT-RWM</i> , b) image couleur fusionnée par <i>GLP-CBD</i> , c) courbe en trait plein, transect de l'image Pan correspondant à l'extrait de la ligne matérialisée par une flèche rouge dans les figures a et b, en pointillés, transect de la modalité rouge obtenue par <i>ATWT-RWM</i> , d) idem pour la modalité PIR fusionnée par <i>ATWT-RWM</i> , e) idem pour la modalité rouge fusionnée avec <i>GLP-CBD</i> , f) idem pour la modalité PIR, fusionnée avec <i>GLP-CBD</i> .	113
6.7	Image Ikonos, échangeur autoroutier, Fredericton (Canada). a) plan des coefficients d'ondelettes situé entre les résolutions spatiales 4 et 8 m pour la modalité Pan, b) idem pour la modalité rouge, c) transect Pan/modalité rouge, d) idem pour la modalité PIR.	114
6.8	Image Ikonos, échangeur autoroutier, Fredericton (Canada). a) plan d'approximation Pan à 4 m, b) modalité rouge originale, c) transect Pan/modalité rouge, d) idem pour la modalité PIR.	115
6.9	Image Quickbird, extraits de la zone périurbaine de Fredericton (Canada). a) Pan,	117

	0.7 m, b) modalité bleue suréchantillonnée par un opérateur bicubique, 2.8 m, c) transects de ces deux images correspondant aux encadrés rouges. Copyright Digital Globe 2002.	
6.10	a) Produit fusionné par <i>ATWT-M3</i> , b) produit fusionné par <i>ATWT-RWM</i> , c) et d) en trait plein, le transect de la modalité Pan, et en pointillés, les transects respectifs des deux produits fusionnés.	118
6.11	Image Quickbird, bord d'un bâtiment de Fredericton (Canada). a) extrait de référence provenant d'une image Pan, b) même image dégradée par <i>ATWT</i> , puis reconstruite avec ses deux plans de coefficients d'ondelettes de plus haute résolution spatiale multipliés par un facteur 2, c) idem mais avec l'algorithme <i>GLP</i> . Copyright Digital Globe 2002.	119
6.12	Image Quickbird, bord d'un bâtiment de Fredericton (Canada). a) extrait de référence provenant d'une image Pan, b) même image dégradée par <i>ATWT</i> , puis reconstruite avec ses deux plans de coefficients d'ondelettes de plus haute résolution spatiale divisés par un facteur 2, c) idem mais avec l'algorithme <i>GLP</i> . Copyright Digital Globe 2002.	120
6.13	Image Pléiades, Toulouse. a) produit fusionné par <i>ATWT-M3</i> , b) par <i>FONC</i> , c) par <i>ATWT-M3-FONC</i> , d) par <i>ATWT-SharpendedM3</i> , e) modalités MS originales. Copyright CNES 2000.	126
6.14	Image Ikonos, Hasselt (Belgique). a) produit fusionné par <i>ATWT-M3</i> , b) par <i>FONC</i> , c) par <i>ATWT-M3-FONC</i> , d) par <i>ATWT-SharpendedM3</i> , e) Pan, f) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Space Imaging 2002.	127
6.15	Image Ikonos, abords de Fredericton (Canada). a) produit fusionné par <i>ATWT-M3</i> , b) par <i>FONC</i> , c) par <i>ATWT-M3-FONC</i> , d) par <i>ATWT-sharpenedM3</i> , e) Pan, f) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Space Imaging 2002.	128
6.16	Image Quickbird, centre ville de Fredericton (Canada). a) produit fusionné par <i>ATWT-M3</i> , b) par <i>FONC</i> , c) par <i>ATWT-M3-FONC</i> , d) par <i>ATWT-sharpenedM3</i> , e) Pan, f) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Digital Globe 2002.	130
6.17	Image Ikonos, Toulouse. a) produit fusionné par <i>ATWT-M3</i> , b) par <i>FONC</i> , c) par <i>ATWT-M3-FONC</i> , d) par <i>ATWT-sharpenedM3</i> , e) Pan, f) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Space Imaging 2002.	131
6.18	Image Quickbird, extrait de la périphérie de Starckville (USA). a) composition colorée obtenue par la méthode <i>FONC</i> , b) extrait Pan correspondant, c) modalités MS originales suréchantillonnées par un opérateur bicubique. Copyright Digital Globe 2003.	132
6.19	Image Pléiades, Toulouse. a) composition colorée obtenue par la méthode <i>FONC</i> , b) extrait Pan correspondant, c) modalités MS de référence à haute résolution spatiale (40 cm). Copyright CNES 2000.	134
6.20	Image Ikonos, Fredericton (USA). a) composition colorée obtenue par la méthode <i>FONC</i> , b) extrait Pan correspondant, c) modalités MS originales	133

suréchantillonnées par un opérateur bicubique. Copyright Space Imaging 2002.

6.21	a) Composition colorée qui reprend l'exemple Pléiades au-dessus de Toulouse, issue de la fusion par <i>ATWT-EnhancedRWM</i> , b) extrait Quickbird au-dessus de Fredericton, même méthode de fusion.	134
6.22	a) Composition colorée qui reprend l'exemple Ikonos au-dessus de Hasselt (Belgique), issue de la fusion par <i>ATWT-EnhancedRWM</i> , b) extrait Ikonos au-dessus d'un échangeur de Fredericton (Canada), même méthode de fusion.	135
6.23	Image Pléiades, Toulouse. a) modalité PIR fusionnées par <i>ATWT-M3</i> , b) par <i>FONC</i> , c) par <i>ATWT-M3-FONC</i> , d) par <i>ATWT-SharpenedM3</i> , e) par <i>ATWT-RWM</i> , f) par <i>ATWT-CBD</i> , g) par <i>ATWT-EnhancedRWM</i> , h) modalité PIR originale. Copyright CNES 2000.	140
<hr/>		
7.1	a) Image à résonance magnétique, b) tomographie par émission de positons et c) produit fusionné à partir des images a et b. (Pajares et Cruz 2004)	147
<hr/>		

Liste des tableaux

<i>Indice du tableau</i>	<i>Intitulé du tableau</i>	<i># page</i>
1.1	Exemples de systèmes d'observation de la Terre offrant diverses images à différentes résolutions spatiales.	4
1.2	Récapitulatif des notations.	13
2.1	Evolution de grandeurs statistiques avec le changement d'échelle (valeurs en niveaux de gris).	19
3.1	Les bilans de qualité considérés pour la validation de l'hypothèse d'extrapolation.	37
3.2	Statistiques monomodales pour l'extrait de la figure 3.2.	38
3.3	Statistiques multimodales pour l'extrait de la figure 3.2.	39
3.4	Statistiques monomodales pour l'extrait de la figure 3.3.	40
3.5	Statistiques multimodales pour l'extrait de la figure 3.3.	40
3.6	Intervalles de validité des distances pour chaque étape de vérification de l'hypothèse.	41
3.7	Forme du tableau de résultats pour un BQ donné.	42
3.8	BQ1 = $[cc]$, résultats à peu près équivalents pour BQ2.	42
3.9	BQ3 = $[\sigma_{Rel}, cc]$	42
3.10	BQ4 = $[diffVarRel, \sigma_{Rel}, cc]$	43
3.11	BQ5 = $[SAM]$	43
3.12	BQ6 = $[M(V_{res}), \sigma(V_{res})]$	43
3.13	BQ7 = $[ERGAS]$	43

3.14	BQ8 = [<i>cc</i> , <i>SAM</i>]	44
3.15	BQ9 = [<i>diffVarRel</i> , σ_{Rel} , <i>cc</i> , <i>ERGAS</i> , <i>SAM</i> , <i>M(Vres)</i> , $\sigma(Vres)$]	44
3.16	Résultats pour les quatre extraits de la figure 3.5.	50
3.17	Résultats pour le premier extrait Pléiades. Pour un BQ donné, 1 signifie que l'hypothèse d'extrapolation est vérifiée, et sinon, la valeur de la case est 0.	53
3.18	Variances globales pour l'extrait 1 pour chaque approximation de l'image Pan. Les deux colonnes de droites correspondent à la différence en variance entre résolutions spatiales successives, relative à la variance de l'approximation à l'échelle plus fine.	53
3.19	Résultats pour l'hypothèse d'extrapolation de l'extrait 2.	54
3.20	Variances globales et différences en variance entre résolutions spatiales successives, relatives à la variance de l'approximation à l'échelle plus fine, pour l'extrait 2.	54
3.21	Nouvelles bornes d'acceptabilité et rappel des valeurs idéales pour chaque distance.	55

4.1	Vérification des valeurs des coefficients de la droite pour le contour considéré – a) valeurs calculées par l'algorithme, b) valeurs obtenues à la main.	64
4.2	Résultats obtenus pour cinq autres contours.	64
4.3	Résultats de la routine d'estimation de la <i>FTM</i> pour la modalité Pan du satellite Ikonos.	68
4.4	Résultats pour la modalité bleue Ikonos.	68
4.5	<i>FTM(f_N)</i> pour les différentes modalités du satellite Ikonos (Dial <i>et al.</i> 2003).	69

6.1	Récapitulatif des images pour la seconde campagne d'expérimentation.	124
6.2	Résultats pour la vérification de la propriété de cohérence, extrait Pléiades.	138
6.3	Résultats pour la vérification de la propriété de synthèse, extrait Pléiades.	142

ANNEXES

1 Interface graphique sous ENVI et codes IDL

Les codes développés sous IDL ont été complétés par une interface avec l'outil de visualisation et de traitements d'images ENVI associé à IDL. Nous avons rajouté une fonctionnalité supplémentaire à la barre d'outils ENVI : *fusion*.

Lorsque l'on sélectionne ce bouton, comme montré sur la figure A1, deux sous-menus apparaissent. Le premier concerne la fusion de type ARSIS, et le second l'évaluation de la qualité de produits fusionnés. Ces deux modules sont totalement indépendants.

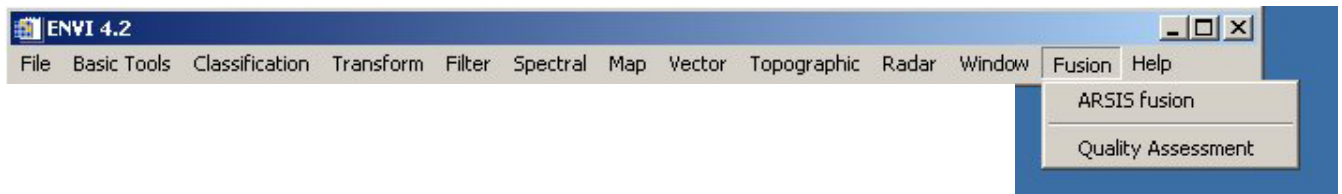


Figure A1 : nouveau menu ENVI.

1.1 Sous-menu 'ARSIS fusion'

Une fois ce sous-menu sélectionné, une fenêtre s'affiche (figure A2). L'utilisateur doit alors rentrer quelques caractéristiques des modalités d'entrée, à savoir :

- ◆ le nom du satellite,
- ◆ le nom de la ville ou de la région représentée sur l'image,
- ◆ la résolution spatiale de la modalité Pan en mètres,
- ◆ celle des modalités MS en mètres,
- ◆ la dynamique des images,
- ◆ le nombre de modalités MS,
- ◆ le *MSM*,
- ◆ le couple *IMM-HRIMM*,
- ◆ la couleur des modalités MS.

Enfin, l'utilisateur choisit de placer les modalités fusionnées soit en mémoire pour pouvoir les visualiser directement sous ENVI, soit directement dans un de ses répertoires. Les deux premières chaînes de caractères (satellite et ville) et les couleurs des modalités MS ne servent qu'à préciser le nom des images fusionnées créées. Ces images sont en format TIFF et possèdent la dynamique des images d'origine. Les trois derniers couples *IMM-HRIMM* correspondent aux nouvelles implantations *ARSIS* développées durant la thèse.

Ensuite, les modalités à fusionner doivent être sélectionnées. Elles doivent auparavant avoir été ouvertes sous ENVI, c'est-à-dire être présentes dans la fenêtre *available bands list*.

ARSIS fusion: Enter input parameters

Satellite name

City name

Resolution of the high resolution image (Pan) (m)

Resolution of the low resolution image (MS) (m)

Dynamic of images (ex: 2048)

Number of MS modalities

☐ ATWT
 ☐ Daub4
 ☐ GLP

☐ M1
 ☐ M2
 ☐ M3
 ☐ R'WM
 ☐ CBD
 ☐ SharpenedM3
 ☐ EnhancedR'WM
 ☐ WindowAdjustR'WM

Select colors of MS channels

☐ [1] pir
☐ [2] red
☐ [3] grn
☐ [4] blu

Number of items selected:

Output Result to ☒ File ☐ Memory

Enter Output Filename

Figure A2 : sous-menu *ARSIS fusion*.

1.2 Sous-menu ‘Quality assessment’

La première étape est d’ouvrir les modalités fusionnées et celles de référence pour les rendre disponibles dans la fenêtre *available bands list*. Ensuite, l’utilisateur sélectionne le sous-menu *Quality assessment*, ce qui affiche la fenêtre placée en figure A3. Les informations nécessaires au fonctionnement de cette routine sont :

- ◆ la résolution des images MS,
- ◆ le ratio de résolution,
- ◆ la dynamique des images,
- ◆ le bord en nombre de pixels à exclure lors des calculs,
- ◆ les couleurs des modalités MS,
- ◆ enfin le type de sortie : mémoire (sous ENVI), ou dans un répertoire.

Quality Assessment procedure

Resolution of the low resolution image (m)

Enter ratio of synthesis

Dynamics of images

Discarded edge for statistics (width in pixels)

Select items from the list:

☐ [1] pir

☐ [2] red

☐ [3] grn

☐ [4] blu

Number of items selected:

Select All Clear

OK Cancel

Figure A3 : sous-menu *Quality Assessment*.

Un exemple de sortie de la routine *Quality assessment* est placée en figure A4 : trois modalités MS ont été traitées, correspondant aux modalités PIR, rouge et verte. Les distances monomodales et multimodales affichées représentent le bilan de qualité que nous avons sélectionné suite au chapitre 2 de ce document de thèse. La dernière ligne de la fenêtre informe l'utilisateur que trois fichiers au format texte ont été sauvegardés dans le fichier de sortie défini par défaut dans les préférences ENVI. Le premier des trois fichiers contient les distances visualisées à l'écran, c'est-à-dire le bilan de qualité que nous recommandons. Les deux autres contiennent respectivement un plus grand nombre de distances respectivement monomodales et multimodales.

```

Quality Budget
File
Monomodal statistics:

Modality                pir
Relative bias (ideal=0)  0.41105876
Relative difference variance (ideal=0)  11.808754
Relative standard deviation (ideal=0)  11.883727
Correlation (ideal=1)    0.968932
Universal Quality Index Q (ideal=1)    0.96701447
Correlation between high frequencies (ideal=1)  0.914338

Modality                red
Relative bias (ideal=0)  0.73261595
Relative difference variance (ideal=0)  9.7139296
Relative standard deviation (ideal=0)  5.2652001
Correlation (ideal=1)    0.997392
Universal Quality Index Q (ideal=1)    0.99606420
Correlation between high frequencies (ideal=1)  0.996120

Modality                grn
Relative bias (ideal=0)  0.85195033
Relative difference variance (ideal=0)  11.762497
Relative standard deviation (ideal=0)  6.5705076
Correlation (ideal=1)    0.995077
Universal Quality Index Q (ideal=1)    0.99309573
Correlation between high frequencies (ideal=1)  0.995053
-----

Multimodal statistics:

ERGAS (ideal=0)          2.1092426
Mean SAM (ideal=0)       1.9887960
Relative bias in norms of spectra (ideal=0)  6.3174768
Standard deviation in norms of spectra (ideal=0)  97.875239
Mean value of the resultant vectors (ideal=0)  76.204613
Standard deviation of the resultant vectors (ideal=0)  92.106682

Statistics are saved in the directory specified in your output default directory

```

Figure A4 : sortie de la routine *Quality assessment*.

2 Description et objectifs de la première campagne d'expérimentation

Description de l'expérience à mener du 23 au 26 janvier 2006

Claire Thomas

Document à remettre aux interprètes

Cette première expérimentation consiste à réunir un groupe de quatre interprètes durant quatre jours, et ce à plein temps, pour juger de la qualité de différents produits issus de divers algorithmes de fusion d'images satellite. L'objet de la fusion est de synthétiser des images couleur de haute résolution spatiale à partir d'un ensemble d'images couleur de moins bonne résolution et d'une image Panchromatique de haute résolution spatiale.

Les objectifs de cette expérimentation sont de trois types :

- ♦ l'objectif principal est l'identification des artefacts visuels issus des six différents algorithmes de fusion testés, vis-à-vis des différents types d'images et pour une application donnée. Ces artefacts peuvent être de deux types : géométriques ou liés à la couleur,
- ♦ ces observations mènent, dans une seconde étape, à une comparaison relative des méthodes,
- ♦ enfin, à terme, nous explorerons la possibilité de proposer une ou plusieurs distances, formant ainsi ce que nous appellerons un "bilan de qualité", qui permettent de rendre compte de la qualité observée par les interprètes.

2.1 Description du Corpus

2.1.1 Types d'interprètes images

Les interprètes qui vont mener cette étude travaillent dans trois grands domaines d'application de l'imagerie satellite, qui sont la détection, la reconnaissance et l'identification d'objets, la cartographie, et enfin la préparation de mission. Par préparation de mission, on entend l'analyse du terrain et des infrastructures par satellite pour l'envoi de moyens en hommes ou véhicules sur le terrain.

2.1.2 Définition d'un cas

Un cas est constitué de :

- ♦ une image Panchromatique (Pan) de haute résolution spatiale,
 - ♦ un ensemble d'images en couleurs naturelles, appelées « vraies couleurs » de résolution moins bonne,
 - ♦ et six produits de fusion en vraies couleurs de haute résolution,
- ou, dans d'autres cas, de
- ♦ un ensemble d'images en couleurs non naturelles, appelées « fausses couleurs » de résolution moins bonne,
 - ♦ et six produits de fusion en fausses couleurs de haute résolution.

Les images couvrent la même région géographique, sont simulées comme acquises par le même satellite et de manière simultanée.

2.1.3 Dix cas (K) à traiter

Voici la synthèse des cas qui sont à notre disposition :

<i>Cas</i>	<i>Images disponibles</i>
K1 : SPOT5 - Toulouse – aéroport - image en fausses couleurs	<p>K1_Pan (5 m), image Pan originale K1orig (10 m), image multispectrale originale</p> <p>K1_M1 (5 m), image en fausses couleurs, méthode M1 K1_M2 (5 m), image en fausses couleurs, méthode M2 K1_M3 (5 m), image en fausses couleurs, méthode M3 K1_M4 (5 m), image en fausses couleurs, méthode M4 K1_M5 (5 m), image en fausses couleurs, méthode M5 K1_M6 (5 m), image en fausses couleurs, méthode M6</p>
K2 : Ikonos - Fredrickton - centre ville – image en vraies couleurs	<p>K2_Pan (1 m), image Pan originale K2orig (4 m), image multispectrale originale</p> <p>K2_M1 (1 m), image en vraies couleurs, méthode M1 K2_M2 (1 m), image en vraies couleurs, méthode M2 K2_M3 (1 m), image en vraies couleurs, méthode M3 K2_M4 (1 m), image en vraies couleurs, méthode M4 K2_M5 (1 m), image en vraies couleurs, méthode M5 K2_M6 (1 m), image en vraies couleurs, méthode M6</p>
K3 : Ikonos - Toulouse - gare– image en vraies couleurs	<p>K3_Pan (1 m), image Pan originale K3orig (4 m), image multispectrale originale</p> <p>K3_M1 (1 m), image en vraies couleurs, méthode M1 K3_M2 (1 m), image en vraies couleurs, méthode M2 K3_M3 (1 m), image en vraies couleurs, méthode M3 K3_M4 (1 m), image en vraies couleurs, méthode M4 K3_M5 (1 m), image en vraies couleurs, méthode M5 K3_M6 (1 m), image en vraies couleurs, méthode M6</p>
K4 : Ikonos - Toulouse – aéroport – image en vraies couleurs	<p>K4_Pan (1 m), image Pan originale K4orig (4 m), image multispectrale originale</p> <p>K4_M1 (1 m), image en vraies couleurs, méthode M1 K4_M2 (1 m), image en vraies couleurs, méthode M2 K4_M3 (1 m), image en vraies couleurs, méthode M3 K4_M4 (1 m), image en vraies couleurs, méthode M4 K4_M5 (1 m), image en vraies couleurs, méthode M5 K4_M6 (1 m), image en vraies couleurs, méthode M6</p>
K5 : Ikonos – Toulouse – aéroport – image en vraies couleurs	<p>K5-Pan (1 m), image Pan originale K5orig (4 m), image multispectrale originale</p> <p>K5_M1 (1 m), image en vraies couleurs, méthode M1 K5_M2 (1 m), image en vraies couleurs, méthode M2 K5_M3 (1 m), image en vraies couleurs, méthode M3 K5_M4 (1 m), image en vraies couleurs, méthode M4 K5_M5 (1 m), image en vraies couleurs, méthode M5 K5_M6 (1 m), image en vraies couleurs, méthode M6</p>
K6 : Ikonos - Hasselt – zone portuaire - image en vraies couleurs	<p>K6_Pan (1 m), image Pan originale K6orig (4 m), image multispectrale originale</p> <p>K6_M1 (1 m), image en vraies couleurs, méthode M1</p>

	K6_M2 (1 m), image en vraies couleurs, méthode M2 K6_M3 (1 m), image en vraies couleurs, méthode M3 K6_M4 (1 m), image en vraies couleurs, méthode M4 K6_M5 (1 m), image en vraies couleurs, méthode M5 K6_M6 (1 m), image en vraies couleurs, méthode M6
K7 : Quickbird - Fredrickton – centre ville – image en vraies couleurs	K7_Pan (0.7 m), image Pan originale K7orig (2.8 m), image multispectrale originale K7_M1 (0.7 m), image en vraies couleurs, méthode M1 K7_M2 (0.7 m), image en vraies couleurs, méthode M2 K7_M3 (0.7 m), image en vraies couleurs, méthode M3 K7_M4 (0.7 m), image en vraies couleurs, méthode M4 K7_M5 (0.7 m), image en vraies couleurs, méthode M5 K7_M6 (0.7 m), image en vraies couleurs, méthode M6
K8 : Quickbird - Fredrickton - exploitations agricoles – image en fausses couleurs	K8_Pan (0.7 m), image Pan originale K8orig (2.8 m), image multispectrale originale K8_M1 (0.7 m), image en fausses couleurs, méthode M1 K8_M2 (0.7 m), image en fausses couleurs, méthode M2 K8_M3 (0.7 m), image en fausses couleurs, méthode M3 K8_M4 (0.7 m), image en fausses couleurs, méthode M4 K8_M5 (0.7 m), image en fausses couleurs, méthode M5 K8_M6 (0.7 m), image en fausses couleurs, méthode M6
K9 : Quickbird - Fredrickton - exploitations agricoles – images en vraies couleurs	K9_Pan (0.7 m), image Pan originale K9orig (2.8 m), image multispectrale originale K9_M1 (0.7 m), image en vraies couleurs, méthode M1 K9_M2 (0.7 m), image en vraies couleurs, méthode M2 K9_M3 (0.7 m), image en vraies couleurs, méthode M3 K9_M4 (0.7 m), image en vraies couleurs, méthode M4 K9_M5 (0.7 m), image en vraies couleurs, méthode M5 K9_M6 (0.7 m), image en vraies couleurs, méthode M6
K10 : Simulation Pléiades 40cm - Toulouse – images en vraies couleurs	K10_Pan (0.4 m), image Pan originale K10orig (1,6 m), image multispectrale originale K10_M1 (0.4 m), image en vraies couleurs, méthode M1 K10_M2 (0.4 m), image en vraies couleurs, méthode M2 K10_M3 (0.4 m), image en vraies couleurs, méthode M3 K10_M4 (0.4 m), image en vraies couleurs, méthode M4 K10_M5 (0.4 m), image en vraies couleurs, méthode M5 K10_M6 (0.4 m), image en vraies couleurs, méthode M6

2.2 Méthodes de fusion testées (pour information)

Les méthodes de fusion testées sont au nombre de six. Nous les indiquerons de 1 à 6 de la manière suivante : M1, M2... à M6. Ces méthodes sont :

1. M1 = transformée en ondelettes non-décimée + modèle global M3 basé sur les détails
2. M2 = transformée en ondelettes non-décimée + modèle local RWM basé sur les détails
3. M3 = pyramides Laplaciennes + modèle basé sur les contextes
4. M4 = pyramides Laplaciennes + M3
5. M5 = Mallat + RWM
6. M6 = minimisation d'une fonctionnelle d'énergie

2.3 Un mode opératoire en 2 étapes

Les images sont au format TIFF. Le logiciel de visualisation des images que vous allez utiliser est le logiciel ENVI. Si vous effectuez des post-traitements (déconvolution, rehaussement des contours, du contraste) pour aider à la visualisation de certaines régions de l'image, décrivez précisément votre démarche, et appliquez les mêmes opérations pour toutes les images d'un même cas. N'hésitez pas non plus à ajouter des commentaires supplémentaires à ceux que nous vous proposons.

2.3.1 Etape 1 : évaluation de chaque méthode

Cette étape est à effectuer pour chaque méthode de fusion pour chacun des cas à traiter.

2.3.1.1 Impression générale de l'image fusionnée

Affichez uniquement l'image fusionnée à l'écran. Répondez aux questions qui suivent.

Q1. Aspect géométrique général :

Regardez l'image fusionnée : vous paraît-elle floue, peu contrastée, voilée ou avec un effet de brume ? Ou avez-vous l'impression que l'image présente des contours trop prononcés, trop de piqué ou encore une surtexturation générale ?

Cochez la case entre 'trop de flou' à 'trop de piqué' :

trop de flou		un peu flou		normale		un peu de piqué		trop de piqué	
--------------	--	-------------	--	---------	--	-----------------	--	---------------	--

En commentaires, indiquez si une zone présente en particulier les caractéristiques que vous avez cochées dans le tableau, ou si l'image cumule plusieurs défauts.

Q2. Perception des couleurs :

Les couleurs sont-elles bien perceptibles ? Ont-elles tendance à baver sur les pixels voisins ? Cochez la case correspondante :

très peu perceptibles		peu perceptibles		moyennement perceptibles		bien perceptibles		très bien perceptibles	
-----------------------	--	------------------	--	--------------------------	--	-------------------	--	------------------------	--

En commentaires, indiquez si la contamination spectrale est générale sur l'ensemble de l'image, si une direction est privilégiée ou encore si une couleur a plus tendance à baver qu'une autre. Précisez aussi si le bavé prend la forme d'un rebond, d'une irisation ou d'un halo.

Q3. Perception des couleurs sur une région ou des objets particuliers :

Y a-t-il une région de l'image ou des objets pour lesquels la perception des couleurs est plus mauvaises que pour l'ensemble de l'image ? Cochez :

oui		non	
-----	--	-----	--

En commentaires, précisez si les couleurs bavent de manière générale sur l'ensemble de la région considérée, si une direction est privilégiée ou encore si une couleur a plus tendance à baver qu'une autre. Précisez aussi si le bavé prend la forme d'un rebond, d'une irisation ou d'un halo.

2.3.1.2 Comparaison avec l'image Pan

Affichez l'image fusionnée en parallèle à l'image Pan. Répondez aux questions suivantes :

Q4. Détection de contours linéaires ou des surfaces :

Par rapport à l'image Pan, comment détectez-vous les contours ? Les contours et les surfaces peuvent être des bâtiments, ponts, infrastructures, stades, routes, ports...

Cochez la case correspondante :

beaucoup moins bien		moins bien		pareil		mieux		beaucoup mieux	
---------------------	--	------------	--	--------	--	-------	--	----------------	--

Commentaires : si les contours ne sont pas très bien détectés, indiquez la présence de halo, de rebond, de bavé, débordement de couleur...

Q5. Perception des petits objets

Par rapport à l'image Pan, percevez-vous les petits objets ? Les petits objets peuvent être des voitures, arbres, sillons proches, haies, alignements d'arbres.

Cochez la case correspondante :

beaucoup moins bien		moins bien		pareil		mieux		beaucoup mieux	
---------------------	--	------------	--	--------	--	-------	--	----------------	--

En commentaires, indiquez éventuellement s'il s'agit de perte de détection de ces petits objets ou d'agrégation.

Q6. Perception des objets dans les ombres

Comment percevez-vous les objets dans les ombres ?

beaucoup moins bien		moins bien		pareil		mieux		beaucoup mieux	
---------------------	--	------------	--	--------	--	-------	--	----------------	--

En commentaires, précisez si des artefacts sont apparus dans les ombres à cause du processus de fusion.

2.3.1.3 Comparaison avec l'image originale à basse résolution

Affichez l'image fusionnée avec l'image originale sur un même écran. Répondez aux questions suivantes :

Q7. Détection de contours linéaires ou des surfaces

Même question que Q4 mais en comparant l'image fusionnée avec l'image originale.

Q8. Perception des petits objets

Même question que Q5 mais en comparant l'image fusionnée avec l'image originale.

Q9. Perception des objets dans les ombres

Même question que Q6 mais en comparant l'image fusionnée avec l'image originale.

Q10. Similarité globale des couleurs

Les couleurs observées dans l'image fusionnée sont-elles similaires à celles de l'image couleur d'origine ? Cochez la case correspondante :

éloignées		peu éloignées		identiques	
-----------	--	---------------	--	------------	--

En commentaires, précisez si une couleur particulière a été favorisée ou défavorisée.

Q11. Similarité des couleurs sur une région particulière

Existe-t-il une région de l'image, de taille supérieure à une vingtaine de pixels de l'image fusionnée, qui présente des couleurs différentes par rapport à l'image originale ? Cochez la réponse correspondante :

oui		non	
-----	--	-----	--

En commentaires, si la réponse est positive, précisez la zone de l'image et la couleur privilégiée, s'il s'agit d'inversion de couleurs, ...

2.3.1.4 Acceptabilité du produit fusionné

Affichez à l'écran, en plus de l'image originale et de l'image fusionnée en cours de traitement, l'image Pan. Répondez ensuite aux questions suivantes :

Q12. Apport de la fusion

La fusion apporte-t-elle réellement quelque chose de plus par rapport aux ensembles originaux ? Les apports de la fusion sont-ils : Cochez la case correspondante :

très peu perceptibles		peu perceptibles		moyennement perceptibles		bien perceptibles		très bien perceptibles	
-----------------------	--	------------------	--	--------------------------	--	-------------------	--	------------------------	--

Commentaires :

Q13. Fausses interprétations

Est-ce que la fusion fait apparaître de fausses interprétations pour une application donnée ? Les fausses interprétations sont-elles : Cochez la case correspondante :

nombreuses		peu nombreuses		absentes	
------------	--	----------------	--	----------	--

Commentaires :

Q14. Fiabilité des résultats pour une application donnée

Peut-on se satisfaire **uniquement** de l'image fusionnée pour satisfaire une demande particulière ?

Cochez la case correspondante :

oui		non	
-----	--	-----	--

Commentaires :

Q15. Acceptabilité

Enfin, le produit fusionné est-il acceptable ? Cochez la case correspondante :

oui		non	
-----	--	-----	--

Commentaires :

2.3.2 Etape 2 : classement relatif des méthodes de fusion

Q16. Classement des méthodes acceptées

Affichez les produits fusionnés qui ont été acceptés à la question 15 et classez ces produits par ordre croissant de qualité.

NB : pour alléger le texte, nous avons volontairement omis les annexes. Annexe A8

3 Mode opératoire de la seconde campagne d'expérimentation

Description de l'expérience à mener du 12 au 16 juin 2006

Claire Thomas

Document à remettre aux interprètes

3.1 Rappel sur la première campagne

La première campagne a réuni quatre interprètes d'images pour juger de la qualité de cinq algorithmes de fusion d'images. Nous rappelons que l'objet de la fusion est de synthétiser des images couleur de haute résolution spatiale à partir d'un ensemble d'images couleur de moins bonne résolution et d'une image Panchromatique de haute résolution spatiale. Les méthodes de fusion testées appartenaient à la même catégorie de méthodes, le concept ARSIS (amélioration de la résolution spatiale par injection de structure). Leur avantage était de donner de bons bilans de qualité, c'est-à-dire de bons résultats statistiques. Cependant, des études précédentes avaient soulevé le problème de la qualité visuelle des produits de fusion.

Les objectifs étaient l'identification des artefacts visuels et la classification relative des méthodes. Les interprètes d'images avaient été regroupés par groupe de deux, avec deux écrans par groupe pour avoir la possibilité de visualiser l'ensemble des images fusionnées de manière simultanée. Pour guider les expérimentateurs, nous leur avons proposé de remplir un questionnaire PhP en ligne. Dix images avaient été traitées par les algorithmes. Ils avaient du rendre un rapport sur leurs observations un mois plus tard et nous faire une présentation orale de leurs résultats sous forme d'un bilan.

3.2 Description et objectifs de la nouvelle expérimentation

3.2.1 Cas traités

Cette seconde expérimentation s'enrichit des remarques et modalités de la campagne précédente. Nous allons réunir à nouveau un groupe de quatre interprètes d'images pendant une semaine à plein temps, pour juger de la qualité de quatre algorithmes de fusion. Parmi ces quatre algorithmes, nous essaierons d'en proposer au moins deux nouveaux.

Comme pour la première campagne, un cas est défini de la manière suivante :

- ◆ une image Panchromatique (Pan) de haute résolution spatiale,
- ◆ un ensemble d'images en couleurs naturelles, appelées « vraies couleurs », de résolution moins bonne,
- ◆ et quatre produits de fusion en vraies couleurs de haute résolution.

Les images couvrent la même région géographique. Elles sont acquises, ou simulées comme acquises, par le même satellite et de manière simultanée. Pour cette nouvelle session, les expérimentateurs nous ont informés de leur souhait de choisir au préalable des extraits présentant des objets d'intérêt dans les images. Pour cela, nous leur avons proposé un ensemble d'images Pan dans lesquels ils ont procédé à des extractions. Nous avons donc ajouté un indice supplémentaire pour

distinguer les différents extraits dans un même cas. La synthèse des sept cas de la campagne 2 est donnée dans le tableau 1 suivant. Par exemple, le troisième extrait du cas numéro cinq (K5) fusionnée par la méthode M3 est appelée K5_M3_3.

<i>Cas (K)</i>	<i>Images disponibles</i>
K1 : Ikonos - Hasselt (Belgique) Nombre total d'extraits : 7	K1_Pan_i (1 m), image Pan originale K1orig_i (4 m), image multispectrale originale (MS) K1_M1_i (1 m), méthode M1 K1_M2_i (1 m), méthode M2 K1_M3_i (1 m), méthode M3 K1_M4_i (1 m), méthode M4
K2 : Ikonos – Fredrickton (Canada) Nombre total d'extraits : 5	K2_Pan_i (1 m), image Pan originale K2orig_i (4 m), image multispectrale originale (MS) K2_M1_i (1 m), méthode M1 K2_M2_i (1 m), méthode M2 K2_M3_i (1 m), méthode M3 K2_M4_i (1 m), méthode M4
K3 : Quickbird – Fredrickton (Canada) - mêmes extraits que K2, + 1 extrait supplémentaire (zone agricole)	K3_Pan_i (0.7 m), image Pan originale K3orig_i (2.8 m), image multispectrale originale (MS) K3_M1_i (0.7 m), méthode M1 K3_M2_i (0.7 m), méthode M2 K3_M3_i (0.7 m), méthode M3 K3_M4_i (0.7 m), méthode M4
K4 : Quickbird – Starckville (USA) Nombre total d'extraits : 8	K4_Pan_i (0.7 m), image Pan originale K4orig_i (2.8 m), image multispectrale originale (MS) K4_M1_i (0.7 m), méthode M1 K4_M2_i (0.7 m), méthode M2 K4_M3_i (0.7 m), méthode M3 K4_M4_i (0.7 m), méthode M4
K5 : Ikonos – centre ville de Toulouse Nombre total d'extraits : 4	K5_Pan_i (1 m), image Pan originale K5orig_i (4 m), image multispectrale originale (MS) K5_M1_i (1 m), méthode M1 K5_M2_i (1 m), méthode M2 K5_M3_i (1 m), méthode M3 K5_M4_i (1 m), méthode M4
K6 : Quickbird – centre ville de Toulouse Mêmes extraits que K5	K6_Pan_i (0.7 m), image Pan originale K6orig_i (2.8 m), image multispectrale originale (MS) K6_M1_i (0.7 m), méthode M1 K6_M2_i (0.7 m), méthode M2 K6_M3_i (0.7 m), méthode M3 K6_M4_i (0.7 m), méthode M4
K7 : Pélican - centre ville de Toulouse Mêmes extraits que K5	K7_Pan_i (0.4 m), image Pan originale K7orig_i (1,6 m), image multispectrale originale (MS) K7_M1_i (0.4 m), méthode M1 K7_M2_i (0.4 m), méthode M2 K7_M3_i (0.4 m), méthode M3 K7_M4_i (0.4 m), méthode M4

Tableau 1 : récapitulatif des cas de la seconde campagne

NB : le nombre total d'extraits à traiter est de 38. Les images multispectrales sont en vraies couleurs.

Dans le but de rendre cette expérimentation plus objective, nous avons choisi de la faire en aveugle, ce qui signifie que le code M1 à M4 ne désigne pas toujours la même méthode de fusion. Leur nom est assigné de manière aléatoire pour chaque cas.

3.2.2 Catégories d'évaluation

Pour chacune des catégories qui ont été définies, plusieurs éléments devront être vérifiés par les expérimentateurs. Les réponses seront faites au moyen d'un fichier Excel qui est discuté ultérieurement.

3.2.2.1 Catégorie 1 : impression générale de l'extrait

Quelle est votre impression globale de l'extrait : confort visuel, esthétique, aspect naturel ...

3.2.2.2 Catégorie 2 : géométrie et structure

- ◆ Élément 1 : linéaires, réseaux de transports, communication ...
- ◆ Élément 2 : transitions, séparations, frontières, bordures ...
- ◆ Élément 3 : surfaces, bâtiments, ...
- ◆ Élément 4 : contours
- ◆ Élément 5 : ponctuels, petits objets
- ◆ Élément 6 : artefacts (apparition, agrégation, disparition d'objets, fantômes, repliement spectral, aliasing ...)
- ◆ Élément 7 : ombres (restitution géométrique correcte)

Pour qualifier chaque élément, on pourrait utiliser les termes : lisse, flou, bruité, tramé, haché, zébré, halo, bavé, déformé, cisaillé, contracté, rebond, homogène, effet de brume ...

Il s'agira également de vérifier la fidélité géométrique et des structures par rapport à l'image panchromatique haute résolution.

3.2.2.3 Catégorie 3 : radiométrie et couleurs

- ◆ Élément 1 : dynamique, luminosité, contraste, intensité, teinte et saturation par rapport à l'image couleur originale à plus basse résolution (couleurs fades, intenses, riches, saturées, dominantes ...)
- ◆ Élément 2 : artefacts (halo, irisation, rebond, aplatissement, mélange de couleurs, apparition, agrégation, disparition, fantômes ...)
- ◆ Élément 3 : ombres (conservation de l'information, fidélité du contenu radiométrique)

3.2.2.4 Catégorie 4 : qualité image

Floue, bruitée, tramée, surtexturée, piquée, peu ou trop contrastée, voilée, pixellisée, ...

3.2.3 Questionnaire d'évaluation

Chaque groupe d'expérimentateurs dispose d'un fichier Excel. Il y a une feuille Excel par cas. Pour chaque extrait, le même tableau est à remplir. Ce tableau comprend pour chacune des catégories et chacune des méthodes : une note, les éléments inacceptables et des commentaires et observations.

La note va de 1 à 4 :

- 4 : inacceptable,
- 3 : acceptable,
- 2 : bonne,
- 1 : très bonne.

Cette notation a été choisie car elle est identique à la notation du classement final des méthodes.

Dans la case juste en-dessous, mentionner les éléments éventuellement inacceptables à l'aide des mots clés définis au paragraphe 2.2. (exemple : linéaires).

Enfin, la troisième case recueille de manière libre les commentaires et observations.

Au bas de la feuille Excel, se trouve un tableau final.

- La première cellule concerne l'acceptabilité, on y répondra par O (oui) ou N (non) pour chaque méthode.
- La deuxième ligne donne le classement des méthodes acceptables les unes par rapport aux autres (1 pour la meilleure méthode, ...).
- La troisième ligne répond à la question sur l'apport de la fusion : cocher la case qui convient (croix).
- La quatrième permet des commentaires de manière libre.

3.2.4 Mode opératoire

La phase opératoire de la campagne d'expérimentation sera basée sur les modalités suivantes :

Il est rappelé que l'objectif est d'évaluer chaque méthode, de procéder à un classement relatif inter-méthodes et de définir un seuil d'acceptabilité pour l'usage opérationnel de ces algorithmes de fusion.

Le dispositif d'expérimentation déployé comprendra :

- 2 stations d'expérimentation.

La configuration sera la suivante : un groupe de 2 interprètes par station. Chaque groupe devra tester l'intégralité des extraits tests.

La campagne se déroulera sur quatre jours complets : du lundi midi au vendredi 11 heures à raison de 6 heures par jours (de 9 h00 à 17h00 avec une pause déjeuner et quelques pauses).

Un PC sera mis à disposition de chaque groupe d'évaluation pour permettre la rédaction des réponses au fichier Excel et sauvegarder les copies d'écran afin de justifier les notations réalisées. Ce fichier Excel relatif au questionnaire est joint en annexe.

Des fiches réflexes seront fournies afin de permettre l'accès direct aux informations relatives aux sites, aux mots clés, ...

Les extraits sont au format TIFF. Le logiciel de visualisation des images que les interprètes d'images utiliseront est le logiciel ENVI. Selon la disponibilité, les stations seront équipées en bi-écran afin de disposer d'une ergonomie adaptée.

Les éléments doivent aussi être vérifiés dans les ombres (facilité de lecture, perte d'informations, ...).

Si des post-traitements sont effectués sur un extrait pour aider à la visualisation (déconvolution, rehaussement des contours, du contraste), les interprètes d'images devront absolument préciser leur démarche, et **appliquer les mêmes opérations** pour tous les produits fusionnés d'un même extrait.

Pour chaque scène globale, des zones tests ont été présélectionnées lors de l'établissement du protocole.

Pour chaque extrait, tous les produits (Pan, multispectral original, éventuellement référence, produits fusionnés) seront affichés simultanément.

A l'issue de l'expérimentation, le questionnaire « à chaud » sera remis à l'Ecole des Mines, lors du débriefing du vendredi après-midi..

Voici la liste des actions à appliquer pour chaque cas :

- afficher l'un des extraits pour l'image Pan, l'image multispectrale originale et les produits de fusion de manière simultanée,
- appliquer les post-traitements éventuels, sur tous les cas relatifs à un même extrait
- remplir le fichier Excel pour l'extrait considéré,
- traiter ainsi tous les extraits du cas considéré,
- une fois tous les extraits traités, répondre à la question d'acceptabilité pour chaque méthode,
- puis classer les méthodes de fusion qui ont été retenues,
- enfin, répondre à la question sur l'apport de la fusion : « que préférez-vous : un Pan tout seul ? un Pan + MS à basse résolution ? ou seulement le produit de fusion ? »

Moyens logiciels particuliers :

Deux filtres de rehaussement de contours sont mis à disposition des expérimentateurs. Ils sont intitulés « Laplacian_0.5.ker » et « Laplacian_0.25.ker ». Le premier des deux filtres rehausse plus fortement les contrastes de l'image. Le principe est de convoluer ce filtre avec une image, et d'ajouter le résultat de la convolution à l'image elle-même.

Sous ENVI, la convolution avec des filtres prédéfinis se trouve dans filter/convolution/File/restore kernel. La somme de bandes est dans le menu basic tools/band math. Dans « enter an expression », taper B1+B2. Cliquer sur « add to list » et taper entrée. Il suffit alors de sélectionner pour B1 l'image d'entrée, et pour B2 l'image convoluée que vous aurez gardée en mémoire.

3.2.5 Travaux à conduire à l'issue de l'expérimentation

Le rapport d'expérimentation final complet et détaillé sera à fournir par chaque groupe d'évaluation sous forme papier et numérique à l'officier SETIS et SPOTI/DEOR pour le 6 juillet 2006.

Ce rapport sera enrichi d'illustrations afin de justifier les constations effectuées dans le questionnaire « à chaud ».

Une présentation PowerPoint synthétisant les résultats de chaque groupe d'évaluation sera à fournir lors de la réunion bilan planifiée le 14 septembre à l'Ecole des Mines de Paris à Paris.

3.2.6 Les pistes d'applicabilité au niveau des différents domaines d'exploitation opérationnelle des méthodes retenues comme acceptables pourront être soulignées à l'occasion de cette réunion bilan.

4 Récapitulatif sur les différentes tailles de fenêtre pour l'estimation des paramètres pour des méthodes de type ARSIS

Une discussion avec les auteurs de l'article Ranchin *et al.* (2003) nous a permis d'établir que la formule permettant de calculer automatiquement la taille de la fenêtre pour l'estimation des paramètres de l'IMM, est erronée. Rappel :

$$t = 7(2res1)/res0 + 1 \quad \text{Eq. A1}$$

Cette relation était au départ établie que pour l'algorithme non-décimé *ATWT*. Cette formule donne $t = 15$ pour le ratio 2. Or, la taille de fenêtre de 57 par 57 pixels obtenue pour le ratio 4 est bien trop élevée. En pratique, les auteurs utilisent dans leur implantation $t = 21$ pour *ATWT*. De plus, une critique majeure est que cette formule ne s'adapte pas au type de *MSM* utilisé.

En effet, l'algorithme *ATWT* est un algorithme non-décimé, ce qui implique que toutes les images de détails et d'approximation ont la même taille quelque soit l'échelle. Mais une formule générique devrait être capable de tenir compte de la décimation produite par les autres algorithmes. De plus, cette décimation diffère d'un algorithme à l'autre, comme dans le cas de l'analyse multirésolution défini par *Mallat* ou par l'algorithme *GLP*. Le tableau A1 récapitule les tailles successives des images de détails et d'approximation produite par l'application successive de ces trois algorithmes sur une image 512 par 512 pixels.

	<i>Taille des images</i>	<i>res0</i>	<i>res1</i>	<i>res3</i>
<i>ATWT</i>	approximations	512	512	512
	détails	-	512	512
<i>Mallat</i>	approximations	512	256	128
	détails	-	256	128
<i>GLP</i>	approximations	512	256	128
	détails	-	512	256

Tableau A1 : exemple de l'évolution de la taille des images en fonction du *MSM* et de la résolution spatiale. Cas des *MSM* dyadiques.

Les algorithmes qui répondent au formalisme de *Mallat* produisent des images dont la taille est divisée par deux à chaque échelle pour un algorithme dyadique. L'algorithme *GLP* crée des approximations qui suivent la même logique que dans le cas de *Mallat*, mais a la particularité de produire des images de détails ont les dimensions de l'image d'approximation précédente dans la décomposition.

Ainsi, la taille de la fenêtre d'estimation va dépendre de plusieurs paramètres : tout d'abord, de la résolution spatiale. Dans le cas d'un paysage urbain, il sera judicieux de privilégier les petites fenêtres d'estimation, contrairement à un paysage avec moins de hautes fréquences spatiales, tel qu'un désert où une région montagneuse. De plus, au vue de la décomposition produite par chaque algorithme, la taille de la fenêtre d'estimation dépendra aussi de l'*IMM*, puisque certains *IMM* agissent sur les images d'approximation alors que les autres agissent sur les images de détails.

Nous avons récapitulé dans le tableau A2 les tailles de fenêtre rencontrées dans la littérature. Attention ! les algorithmes *ATWT* et *Mallat* sont associés dans la littérature à des *IMM* qui agissent sur les plans de détails, contrairement à *GLP* que les auteurs associent avec l'*IMM CBD* qui cherche à établir une relation linéaire entre les plans d'approximation ! Entre parenthèses, nous donnons d'autres valeurs possibles pour les tailles de fenêtre.

	Ratio 2	Ratio 3	Ratio 4
<i>ATWT</i>	15 (ou 11 ou 13)	19	21 (ou 23)
<i>Mallat</i>	7	9	15 (ou 11 ou 13)
<i>GLP</i>	7 (ou 5)	?	9 (Ikonos)

Tableau A2 : récapitulatif des tailles de fenêtre d'estimation au niveau de l'*IMM*. Ces tailles sont celles rencontrées dans la littérature. D'autres valeurs possibles ont été placées entre parenthèses.

Nous avons trouvé une relation simple permettant d'obtenir t pour les deux premiers algorithmes dans le cas d'un *IMM* agissant sur les plans de détails. Pour l'algorithme *ATWT*,

$$t = 13 + 4(\text{ratio} - 2) \quad \text{Eq. A2}$$

ce qui donne $t = 13, 17$ et 21 respectivement pour les ratios 2, 3 et 4.

Pour *Mallat*,

$$t = 5 + 4(\text{ratio} - 2) \quad \text{Eq. A3}$$

soit 5, 9 et 13 pour chaque ratio.

Cependant, une recherche un peu plus approfondie serait nécessaire afin de tenir compte à la fois de tous les MSM, mais aussi de l'*IMM* et de la dimension des objets de la scène d'intérêt.

5 Méthode **MallatDaub9/7-WindowAdjustRWM**

Cette nouvelle méthode de fusion de type *ARSIS* correspond à une méthode exploratoire, développée en collaboration avec Gabriel Peyré, docteur en mathématiques appliquées de l'école polytechnique. Il s'agissait ici d'étudier différentes formes, tailles et orientations de la fenêtre d'estimation pour maximiser la corrélation locale. Cette méthode nous a permis de mieux comprendre le comportement local des méthodes de fusion exclusivement basées sur la maximisation du coefficient de corrélation et d'en cerner les limites.

Le *MSM* s'appuie sur l'analyse multirésolution de *Mallat* avec l'ondelettes 9/7 de Daubechies (Mallat 1998, pages 248 à 252). Cette décomposition en ondelettes est à coefficients réels, avec 9 coefficients pour le passe-bas et 7 pour le passe-haut ; nous l'appelons *Daub9/7*. Contrairement à l'ondelettes que nous avons appelé *Daub4*, les filtres utilisés pour *Daub9/7* sont symétriques et centrés, ce qui implique qu'aucun décalage n'est introduit à chaque échelle de décomposition.

5.1 *IMM de MallatDaub9/7-WindowAdjustRWM*

Le point de départ de cette implantation est le couple *IMM-HRIMM RWM*, auquel nous avons ôté l'application conditionnelle. L'*IMM* cherche toujours à maximiser la corrélation locale, mais cette fois-ci par un filtrage directionnel, en adaptant la taille et la forme de la fenêtre d'estimation pour chaque pixel de l'image. En effet, l'utilisation de fenêtres carrées permet pas de capter la régularité locale d'une image si celle-ci présente une forme plus allongée dans une direction donnée.

Pour exploiter au mieux la géométrie de l'image, on définit une fonction de pondération bi-directionnelle qui s'allonge dans une direction donnée, et que nous avons choisie de type gaussienne $g(\sigma_1, \sigma_2, \theta)$. Les paramètres σ_1 et σ_2 contrôlent la longueur et la largeur de la gaussienne, et θ définit l'orientation de la fonction par rapport à l'axe des abscisses ($\theta = 0$). Cette fonction de pondération est définie sur une fenêtre de taille carrée de taille t par t pixels. On la désigne par $g_t(\sigma_1, \sigma_2, \theta)$; son intégrale sur la fenêtre vaut 1. Nous avons placé un exemple de fonction gaussienne sur la figure A5, où le centre de la fenêtre qui montre les pixels les plus sombres correspond aux valeurs les plus élevées de la gaussienne qui décroissent en s'éloignant progressivement du centre vers le bord de la fenêtre, pour atteindre des valeurs quasiment nulles sur les bords (en blanc sur la figure).

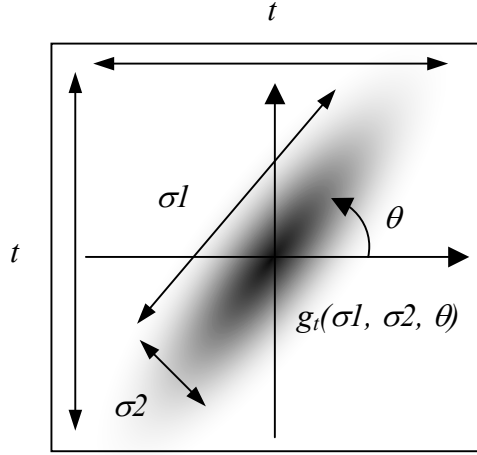


Figure A5 : fenêtre de pondération de taille t par t pixels.

Dans *RWM*, les valeurs de pondération de la fenêtre d'estimation étaient implicitement fixées à 1. Soient $(D_{A,1-2})_t$ et $(D_{Bk,1-2})_t$ les extraits de taille t par t pixels des plans d'ondelettes respectivement pour la modalité Pan et la $k^{\text{ième}}$ modalité MS. On cherche à maximiser la fonction de corrélation *correl* entre ces deux extraits pondérés, par un produit terme à terme, par la gaussienne, en fonction de l'angle θ et des paramètres $\sigma1$ et $\sigma2$, c'est-à-dire entre $[(D_{A,1-2})_t g_t(\sigma1, \sigma2, \theta)]$ et $[(D_{Bk,1-2})_t g_t(\sigma1, \sigma2, \theta)]$:

$$\underset{\theta, \sigma1 \text{ et } \sigma2}{\text{Argmax}}(\text{correl}((D_{A,1-2})_t g_t(\sigma1, \sigma2, \theta), (D_{Bk,1-2})_t g_t(\sigma1, \sigma2, \theta))) \quad \text{Eq. A4}$$

Le produit entre la gaussienne et les extraits de détails est maximal lorsque la direction privilégiée de la fonction de pondération entre en résonance avec le phénomène étudié. Admettons que les deux extraits présentent des structures selon deux directions privilégiées pour $\theta = \pi/4$ et $\theta = 5\pi/3$ (figure A6a). La figure A6b montre la réponse du coefficient de corrélation en fonction de θ , pour un couple $(\sigma1, \sigma2)$ fixé.

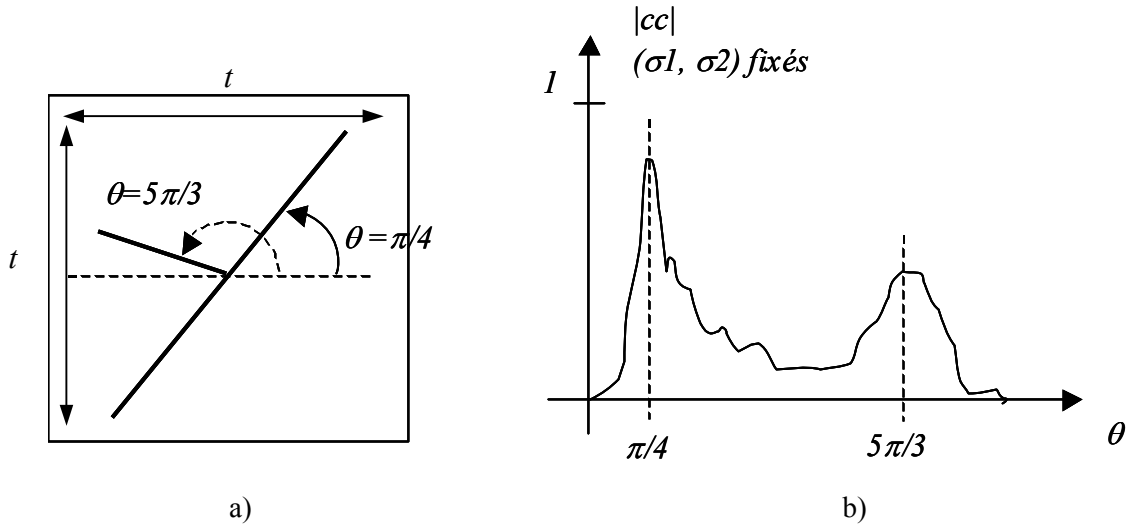


Figure A6 : a) directions privilégiées dans les deux extraits $(D_{A,1-2})_t$ et $(D_{Bk,1-2})_t$, b) résultats obtenus en fonction de θ , avec $(\sigma1, \sigma2)$ fixés.

Si la corrélation est maximisée sans contrainte, on tombe dans le piège où ce sont toujours les plus petites fenêtres qui sont favorisées. En effet, la corrélation se trouve généralement inversement proportionnelle à la taille de la fenêtre. Voilà pourquoi nous avons multiplié le terme à maximiser par l'inverse de la taille de la fenêtre.

Le paramètre a est estimé par la méthode des moindres carrés sur la fenêtre sélectionnée. Le paramètre b est implicitement placé à zéro.

5.2 *HRIMM de MallatDaub9/7-WindowAdjustRWM*

Le *HRIMM* s'écrit :

$$D_{Bk,0-l} = a D_{A,0-l}$$

Eq. A5

5.3 *Résultats*

La figure A7 montre comment la fenêtre d'estimation se déforme en fonction du contenu de l'image. Pour cela, on a superposé à une photographie de visage les gaussiennes déformées, en affichant une gaussienne tous les dix pixels. On remarque que les zones homogènes comme le front ou les joues privilégient les fenêtres isotropes, alors que les contours du visage et des éléments du visage ont fortement orienté la fenêtre. Cette étape du calcul est validée sur cet exemple.



Figure A7 : vérification de l'orientation de la fenêtre en fonction de la géométrie de l'image.

La figure A8 suivante montre les résultats obtenus sur trois des quatre extraits utilisés pour vérifier si les nouvelles méthodes de fusion développées présentaient ou non les artefacts de la première campagne d'expérimentation.



Figure A8 : résultats obtenus par la méthode *MallatDaub9/7-WindowAdjustRWM* sur les extraits : a) la piste d'athlétisme de Hasselt, b) l'échangeur autoroutier de Fredericton et c) bâtiment de Fredericton.

On remarque que les résultats obtenus par cette méthode de fusion sont meilleurs que ceux obtenus par *ATWT-EnhancedRWM*. En effet, on note que globalement, les artefacts colorés et les amas de pixels de couleurs aberrantes sont moins prononcés que sur les extraits traités par la méthode *ATWT-EnhancedRWM*. L'extrait figure A8a montre une bonne cohérence des couleurs avec la composition colorée des modalités MS d'origine, mais montre aussi le défaut de faïençage relevé lors de la première campagne. On remarque toutefois que cet extrait a été mieux synthétisé qu'avec les deux méthodes locales de la première campagne puisque l'on ne retrouve pas l'amas de pixels roses de la piste d'athlétisme et les autres ponctuels colorés. La figure A8b montre encore le défaut d'inversion de contraste. Les voies de circulation et leurs signalisations ont été synthétisées avec la mauvaise radiométrie, ce qui est en accord avec la discussion du chapitre 6 sur le traitement de cet extrait particulier par les approches locales. Les couleurs des routes sont aussi erronées avec une tendance à tirer vers le mauve ou le bleu clair. Les zones de transitions entre le bâti et la végétation ont été

touchées par l'incertitude liée à la valeur de la corrélation entre le Pan et les MS sur ces zones critiques, toutefois, l'intérieur de ces deux types de paysage ont été très bien traités. Enfin, la qualité du dernier extrait est très bonne, avec un bon rendu du contour du bâtiment. Aucun flou n'est venu gommé la transition blanc/noir. Par contre, le véhicule rouge en haut de l'image est complètement floue puisque sa couleur a bavé sur la radiométrie des pixels du bitume.

En conclusion, nous avons placé cette méthode de fusion en annexe car elle ne répondait pas aux exigences que l'on s'était fixées lors de la fin de la première campagne. En effet, cette méthode ne correspond pas à une combinaison entre une approche globale et locale comme nous le recommandions. Par contre, son intérêt est que cela nous a permis de pousser l'approche locale basée exclusivement sur la corrélation locale, dans ses derniers retranchements. Ces quelques exemples montrent que même en adaptant la taille et la pondération de la fenêtre, nous ne sommes pas en mesure de nous affranchir du défaut de type faïençage. Sur les régions où la corrélation est bonne entre le Pan et les MS, c'est-à-dire sur le bâti, la qualité du résultat est remarquable, avec toutefois quelques difficultés à traiter les objets en mouvement comme sur l'exemple du véhicule rouge de la figure A8c. Nous pensons qu'un paramètre supplémentaire doit intervenir pour compléter l'information rapportée par le coefficient de corrélation qui n'est pas capable à lui seul de respecter la régularité des images, comme le ferait une approche exclusivement globale.

FUSION D'IMAGES DE RESOLUTIONS SPATIALES DIFFERENTES

Résumé

Les satellites de type SPOT, Ikonos ou Quickbird fournissent des images possédant des caractéristiques différentes : d'une part des images de basse résolution spectrale et de haute résolution spatiale, et d'autre part, des images de caractéristiques opposées, c'est-à-dire de haute résolution spectrale combinée à une basse résolution spatiale. De nombreuses applications nécessitent de combiner ces hautes résolutions spatiales et spectrales. Ce travail s'insère dans la synthèse d'images multispectrales à haute résolution au moyen de la fusion d'images. La constatation de lacunes dans la littérature à propos de l'évaluation de la qualité, a mené à l'établissement d'un protocole d'évaluation de la qualité, correspondant à une généralisation de deux travaux précédents. Une étude empirique sur l'évolution de la qualité des produits de fusion dans les échelles complète ce protocole. Une catégorisation des nombreuses distances de la littérature est proposée. Nous ajoutons une nouvelle distance pour caractériser la qualité géométrique des produits de fusion basée sur la fonction de transfert de modulation, que nous avons validée grâce à l'imagerie Ikonos et à son application à quelques méthodes de fusion connues. La sélection d'une série de distances sous forme de bilan de qualité est discutée dans le but d'obtenir une idée de la qualité qui soit la plus complète possible.

Puisque nous disposons des outils nécessaires à l'évaluation de la qualité, nous pouvons désormais les exploiter pour le développement de nouvelles méthodes de fusion. Nous avons choisi le concept ARSIS comme cadre de développement car, d'une part, ses différentes implantations donnent généralement de bons bilans de qualité, d'autre part, leur qualité visuelle a été parfois controversée. Une campagne d'évaluation de la qualité visuelle des produits de fusion a permis de répertorier, de classer puis d'expliquer la formation des artefacts de quelques implantations ARSIS. L'étude critique de ces méthodes a mené au développement de trois nouvelles méthodes de fusion qui s'appuient aussi sur des outils extérieurs au concept. Ces nouvelles méthodes sont testées lors d'une seconde campagne d'expérimentation. Une des nouvelles méthodes offre de meilleurs résultats que la méthode de référence choisie. Les expérimentateurs l'ont déclarée acceptable et recommandable pour une utilisation dans le cadre d'une exploitation opérationnelle au sein du ministère de la défense.

Mots clés : fusion d'images, fusion de données, évaluation de la qualité, modalités panchromatique et multispectrales, concept ARSIS, infrarouge, télédétection, analyse multiéchelle et multirésolution, transformée en ondelettes, approche variationnelle, estimation de paramètres, fonction de transfert de modulation

FUSION OF IMAGES OF DIFFERENT SPATIAL RESOLUTIONS

Abstract

Space-borne sensors such as SPOT, Ikonos or Quickbird, provide images with different characteristics : on one hand, images with high spectral resolution but low spatial resolutions, and on the other hand, images with high spatial resolution combined to low spectral resolution. This work deals with the synthesis of multispectral images at high spatial resolution by the mean of image fusion. Weak approaches to quality assessment were found in literature. Accordingly, we design a protocol for quality assessment based on two previous relevant works. An empirical survey about the changes in quality budget within scales completes the description of the protocol. We propose a categorization of distances found in literature, and add a new distance for geometrical quality assessment, based on the Modulation Transfer Function. It was validated on Ikonos actual imagery and synthesized images obtained by well-known fusion methods. The selection of a group of distances to form quality budgets is discussed in order to obtain a complete overview of the quality reached by fused products.

Once the necessary tools of quality assessment built, we focused on the development of new fusion methods. We have chosen the ARSIS concept as framework because of the good quality budgets attained by its different implementations. However, the visual quality of the resulting images is often criticized. A campaign for the assessment of the visual quality of several ARSIS fused products resulted into the identification, the classification and the explanation of several artefacts. This critical study formed the basis for the development of three new methods. They were tested during a second campaign of experimentation. One of the new methods offers better results than that chosen as reference. The image analysts of this second campaign stated that this method was acceptable and recommendable for an operational exploitation within the minister of defence.

Key words : image fusion, data fusion, quality assessment, panchromatic and multispectral imagery, ARSIS concept, infrared, remote sensing, multiscale and multiresolution analysis, wavelet transform, variational approach, parameter estimation, modulation transfer function

Laboratoire d'accueil : Centre Energétique et Procédés - Ecole des Mines de Paris
Rue Claude Daunesse - B.P. 207 - F-06904 Sophia-Antipolis Cedex

Thèse présentée par : THOMAS Claire le : 11 décembre 2006

Discipline : « Informatique Temps Réel, Robotique, Automatique » - Ecole des Mines de Paris
